Using Quantitative Trait Locus Mapping to Understand Arsenite Tolerance in Arabidopsis

Li, Jane

Award date:
2017

Awarding institution:
University of Bath

Using Quantitative Trait Locus Mapping to Understand Arsenite Tolerance in Arabidopsis

Dan Li

A thesis submitted for the degree of Doctor of philosophy

University of Bath

Department of Biology and Biochemistry

September 2016

Signed:

Acknowledgements

I would like to express my gratitude to my supervisor, Prof. Rod Scott, for his valuable support, helpful guidance and immense trust during my PhD study. I would also like to thank Dr. John Beeching for his encouragement and helpful comments throughout this research. My sincere thanks go to Dr. Nazmul Haq. His concern of arsenic contamination brought me to this project.

I am very grateful to Dr. Adrian Brennan, for his patience, motivation, and knowledge. His guidance helped me in all genetic mapping analysis, especially GWAS analysis. Many thanks to Dr. Tony Perry for his insightful comments and encouragement, but also for the hard questions which enchanted me to broaden my knowledge from various perspectives. I also have to thank Dr. Maki Asam for her experience of molecular biology and encouragement in my research.

My sincere thanks also go to my dear friends Dr. Huijia Wang and Dr. Roger Snowden, their friendship means a lot to me. I am grateful to Dr. Xueying Wang, who showed me the vale of using Adobe Illustrator and now I have beautiful scientific figures in my thesis. Special thanks go to my previous lab mate, Dr. Maha Al-Jabri. She was a good example to show me to be strong and independent. And I also thank my new lab mates Kate Petty, Sarah Chordekar and Mais Sweiss for their friendship and the tea we shared at many late nights in the office. Grateful acknowledgements are given to all members and technical staff in the Biology and Biochemistry department for their kind cooperation, as well as the crop-innovations organization and University of Bath for funding this project.

At the end, I want to thank my Mum, Dad and all other members of my family. They always support my decision and dream. This Ph.D study taught me the value of family. It was their endless love enabled me to go this far.

This thesis is dedicated to my parents and grandparents.

Abstract

Indian mustard (*Brassica juncea*) grown on land with elevated arsenic concentration can absorb and accumulate high levels of arsenic. *B. juncea* is an important oil-bearing crop and broadly grown in Bangladesh and West Bengal, where the land is heavily contaminated with arsenic. Arsenic is documented as a group 1 carcinogen and can enter the biosphere through the food chain. With the high population density and a heavy agriculture-based economy in Bangladesh and West Bengal, issues related to food safety must be addressed to solve the problems associated with arsenic accumulation in *B. juncea* seed oil, which is caused by *B. juncea* tolerance to arsenic toxicity. Therefore, it is important to study the variation of arsenic tolerance among different Indian mustard cultivars and to understand the genetic mechanisms of arsenic tolerance in Indian mustard. Such results from the study could help to reduce potential arsenic toxicity in both plants and humans through diet as well as to benefit local agriculture industry in Bangladesh and West Bengal.

In this research, a consistent trend was found between arsenic concentration and biomass accumulation in all *B. juncea* cultivars. When arsenic concentration increases, the biomass accumulation decreases. A positive correlation was identified between arsenic tolerance levels in *B. juncea* roots and arsenic absorption in the seed and seed oil. The observation of these correlations in *B. juncea* suggests that breeding arsenic resistant *B. juncea* cultivars could increase plant yield but may inadvertently lead to an increased arsenic concentrations in the seed and seed oil. However, variations in arsenic accumulation were observed in both arsenic resistant and susceptible cultivars, which provide possibilities to breed high arsenic-tolerant and low arsenic-absorptive strains of *B. juncea*. To study the genetic mechanisms underlying arsenic tolerance variation in Indian mustard, the model plant *Arabidopsis thaliana* was employed. Arsenic tolerance of various *A. thaliana* accessions was measured and arsenic tolerance genes were screened using a genetic mapping approach. Since no consistent trend was found between arsenic concentrations and

arsenic tolerance in *A. thaliana* root and shoot tissues, it suggests that arsenic has a complex impact on those phenotypes examined in *A. thaliana*. To study the mechanisms underlying genetic variation of *A. thaliana* for arsenic tolerance, gene mapping including quantitative trait locus (QTL) mapping and genome-wide association mapping (GWAS) was employed to identify targeted genomic regions and candidate genes.

One recombined inbred population was employed in QTL mapping and a worldwide collection of 39 *A. thaliana* accessions was used in GWAS analysis. QTL mapping analysis and GWAS revealed similar genomic regions and candidate genes. Among all revealed candidate genes, one gene (gene ID: At5G05560) with a known function of abiotic stress resistance was identified, which may be related to plant arsenic tolerance or detoxification. To understand the function of this gene, however, requires further analysis. Other genes associated with arsenic tolerance or detoxification function in the genetic region selected from gene mapping study may also be identified by further analysis. The functional analysis of candidate genes, and the molecular markers selection from revealed genomic regions could be applied to contribute to breeding *B. juncea* strains with higher arsenic tolerance and lower arsenic accumulation.

Abbreviations and Acronyms

ANOVA	Analysis of variance
As	Arsenic
As(III)	Arsenite
As(V)	Arsenate
A. thaliana	*Arabidopsis thaliana*
BGI	Beijing genomics institute
B. juncea	*Brassica juncea*
CIM	Composite interval mapping
CTAB	Cetyltrimethyl ammonium bromide
EDTA	Ethylenediaminetetraacetic acid
EMBL-EBI	EMBL european bioinformatics institute
GWAS	Genome-wide association study
ICP-MS	Inductively coupled plasma mass spectrometry
IM	Interval mapping
LC50	Half lethal concentration
MALDI-TOF-MS	Matrix assisted laser desorption/Ionization-time-of-flight mass spectrometry
NCBI	National center for biotechnology information
QTL	Quantitative trait loci
SNP	Single nucleotide polymorphisms
TAE	Tris-acetate-EDTA
TAIR	The Arabidopsis information resource
ROS	Reactive oxygen species
%RGR	Root growth recovery

Contents

Chapter 1

General Introduction

1.1 Arsenic – its prevalence and impact on human health

Arsenic caused pollution has drawn public attention because of its toxicity to human health and harm in the environment. The trivalent and pentavalent oxidation states have been well studied in the cases of contamination (Irvine et al. 2006). Their common inorganic forms, arsenite and arsenate, are considered the major cause of most contamination (Hughes 2002). The International Agency for Research on Cancer (IARC) listed inorganic As as a group one carcinogen to humans (http://www.iarc.fr/). When a large amount of As is ingested, it can cause damage to internal organs and ultimately lead to cancer or death. Small amounts of As ingestion can result in chronic poisoning. It can also lead to arrhythmia and damage to blood vessels, as well as to red and to white blood cells (Rosen et al. 2011) and skin damage (Figure 1.1; Atkins et al. 2006).

Figure 1.1. Skin lesions from As poisoning (Image adapted from Atkins et al. 2006).

Arsenic is an ubiquitous chemical element ranked the 20th most abundant element in earth's crust (Zhao et al. 2010). It has four oxidation states, -3, 0, +3, and +5. The toxicity of As compounds is largely related to the valence and methylation state of the compound (Irvine et al. 2006). Compounds with a trivalent oxidation state are considered as arsines and metal arsines. Those compounds are unstable and can be easily converted to a pentavalent oxidation state in oxidizing conditions (Zhao et al. 2010). Biochemical studies have concluded that the toxic effect of As is due to an increase in the level of an intracellular reactive oxygen species (ROS) and the

disruption of ATP production by replacing phosphate in several reactions (Kang et al. 2004). By comparing arsenite (As^{III}) and arsenate (As^{V}), researchers have confirmed that As^{III} is 60 times more toxic than As^{V} (Hossain 2006). The toxicity of As^{III} is due to its interference with many metabolic pathways by reacting with critical thiols in proteins, therefore inhibiting their activities (Sheng et al. 2013). Arsenite may lead to interrupted potential metabolic steps therefore causing genotoxicity, altered DNA methylation, oxidative stress, altered cell proliferation, co-carcinogenesis, and tumor promotion (Hughes 2002). It plays a role in the cause of cancer (Kang et al. 2004; Martinez et al. 2011).

Anthropogenic sources of As contamination include the mining industry and As-based insecticides, herbicides, fungicides, algaecides, sheep dips, wood preservatives and dyes. Additionally feed additives and compounds for the eradication of tapeworm in sheep and cattle can introduce As into both terrestrial and aquatic ecosystems (Mahimairaja 2005). Geochemically, soils can be As-rich parent material because As can easily substitute for silicon, aluminum or iron in silicate minerals (Bhumbla and Keefer 1994). Arsenic is also commonly associated with sulphides, e.g. in sulfide ore deposits. Nature events, such as volcanic activities, wind blown particles, sea salt sprays and microbial volatilization release As. Water runoff from As mines also contains high concentration of As (Fitz and Wenzel 2002; Azcus and Nriagu 1990). Groundwater introduces As into soil, and this is the major cause of world As contamination (Frankenberger and Arshad 2002).

1.2 Current global arsenic pollution situation

The World Health Organization (WHO 2001) has regulated 10μg/L (10ppb, 133 nM) as the toxicity concentration threshold for As in drinking water, but the regulation of As content in food has not been established (Jackson and Punshon 2015). Another issue here is that this threshold only relates to the total As concentration and does not consider the very diverse toxicities of As species, thus the total As concentration is

only one way to reflect toxicity threshold and the actual form of As in drinking water provide additional valuable information.

Previous research reported that as many as 140 million people globally are continuously poisoned by As polluted drinking water, especially in developing countries (Fitz and Wenzel 2002) (Figure 1.2).

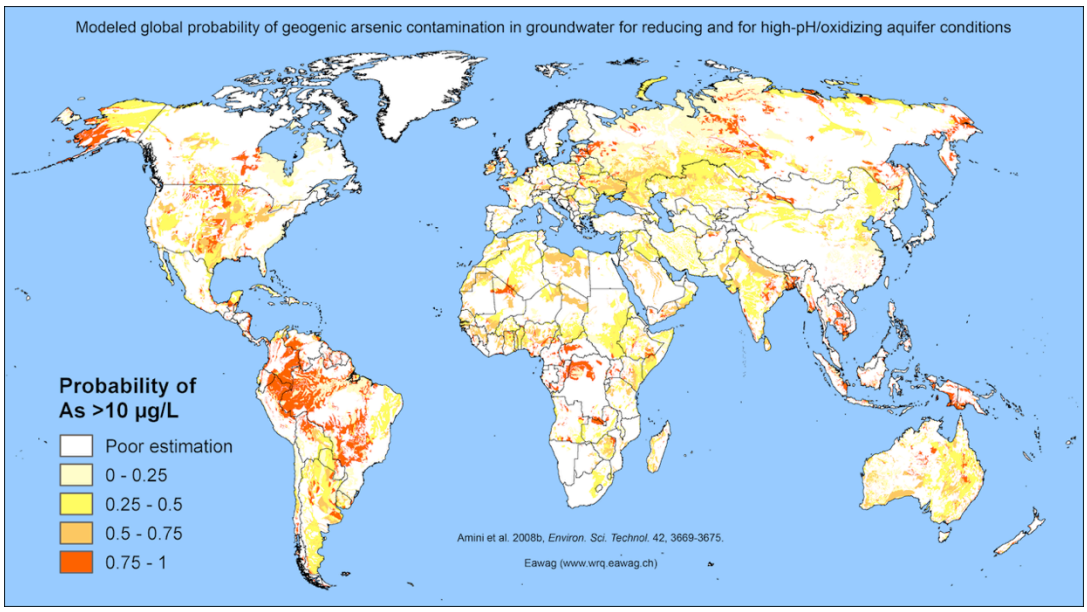

Figure 1.2. Modeled global probability of geogenic As contamination in groundwater conditions (Amini et al. 2008).

Many crops have been reported for their ability to accumulate As in their edible parts (seeds, stems, leaves, roots, fruit, etc.) when grown on As contaminated arable land and/or irrigated with As contaminated water. These crops include rice, wheat, potato, sweet potato, carrot, radish, onion, garlic, cauliflower, lettuce, ginger, tomato, broccoli, cabbage, celery, Chinese cabbage, cucumber, eggplant, spinach, bamboo shoot, snake gourd, ghotkal, taro, green papaya, turmeric powder, beans and green chili (Tripathi et al. 2007). Consuming As polluted food has led to a serious health issue in more than 70 countries, especially in Southeast Asia (Järup 2003).

The most As contaminated area globally is the Bengal Delta (Bangladesh and West Bengal, India) where millions of people depend on As-rich drinking water that is high

in As (Horneman et al. 2004). Groundwater As contamination in Bangladesh is mostly of a non-anthropogenic origin and due to action of the great flood plain of the Bengal Delta. River waters release As from the aquifer sediments into the groundwater; it is then flushed into the surface water, and subsequently transferred to the flood plain where it becomes bioavailable (Horneman et al. 2004).

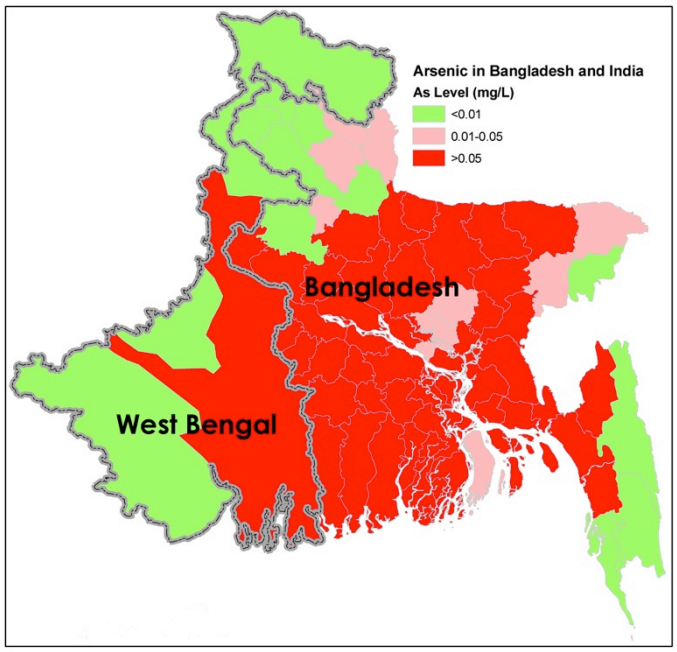

Figure 1.3. Arsenic concentrations in groundwater in Bangladesh and West Bengal (Chowdhury et al. 2000). The red area has an As concentration above 0.05mg/L (0.05ppm); The pink area has an As concentration between 0.01ppm to 0.05ppm; The green area has a concentration less than 0.01ppm.

One research team performed a 10-year survey in 42 districts in Bangladesh with a population of 79.9 million and nine districts in West Bengal with a population of 42.7 million (Chowdhury et al. 2000). The areas covered in the research were 92,106 km^2 in Bangladesh and 38,865 km^2 in West Bengal. They found that about 93% and 77% of samples collected from human hair, nail and urine, contained As that were above the normal/toxic level (Chowdhury et al. 2000). In Bangladesh, more than 2 million people are estimated to face a risk of death from cancer caused by As (WHO 2001).

1.3. Arsenic remediation and solution for Bangladesh

Bangladesh is one of the most densely populated countries in the world, with a population of 161 million people over a comparatively small area of 147, 570km^2 (1092 people/km^2) and a high annual population growth rate of 1.58%. Since much of its soil is contaminated with As, a large number of people, estimated at more than 85 million, are currently suffering from As poisoning (The World Fact book, CIA, 2011). The As concentration in the Bangladesh drinking water is 5 times higher than the WHO recommended safe limit (10 µg/L) (Curry et al. 2000). In some areas of Bangladesh, the reported local groundwater As concentration is as much as 3000ug/mL (Mahimairaja et al. 2005). Arsenic contaminated groundwater is also used to irrigate paddy rice fields. Rice is one of the primary crops in Bangladesh. Around 90% of the people are fed on rice, which provides about 70% of their daily caloric intake. The country has the third biggest total rice growing area in the world and fourth highest rice production (Ahmad et al. 1997; Hossain et al. 2012). The first report of As poisoned rice was widely reported (Meharg and Rahman 2003), and later rice was been found to have a high affinity to bio-accumulate As (Ma et al. 2006). In areas like the As-rich Bengal Delta, it is impractical to stop As from entering the biosphere, owing to its origin from a naturally occurring process. River waters release As from the aquifer sediments into the groundwater, which is then flushed into natural surface water, subsequently when transferred to the flood plain it becomes bioavailable (Horneman et al. 2004). The public health status and economic development of the Bengal Delta is hampered by this naturally occurring As pollution.

With greater public awareness of As poisoning, there has been growing interest in developing guidelines and remediation technologies for mitigating the impacts in As-contaminated ecosystems. Numerous remediation methods have been employed to improve quality of heavy metal contaminated soils, including soil removal and washing, physical stabilisation, and/or the use of chemical amendments; however the cleanup of polluted soils could be very costly (Mahimairaja et al. 2005).

Phytoremediation, which means employing plants to remove heavy metal from soil, is believed to be cost-effective and environmentally friendly (Baker et al. 2000). Using plant in the phytoremediation is called bioremediation, which is a widely accepted method of soil remediation because it is perceived to occur via natural processes (Chibuike and Obiora 2014). Plant cultivation is inexpensive compared to traditional engineering approaches involving intense soil manipulation and minimizing the amounts of secondary waste. Furthermore, this technology creates a minimal environmental disturbance (Tu and Ma 2002). One successful example of using bioremediation was reported by Blaylock et al. (1997). In the study, *B. juncea* was used to enhance accumulation of Pb in plant shoots when grown in Pb-contaminated soil. *B. juncea* can accumulate the concentrations of 1.5% Pb in the shoots from soils containing 600 ppm Pb amended with synthetic chelates such as EDTA, which showed the potential of using *B. juncea* for bioremediation through the application of synthetic chelates to the soil (Blaylock et al. 1997). However, when using chemical such as synthetic chelates, the applied scale of the land is limited, this is one of the shortages of using chemical amended bioremediation.

A successful bioremediation resolution for the Bangladesh As issue must take into account the high population density and the massive scale of contaminated land (Chowdhury et al. 2000). Food and Agriculture Organization and World Bank population estimates the population density was 1, 237 people per sq. km of land area in Bangladesh at 2015 (http://data.worldbank.org/). As the economy of this country is highly based on agriculture, and high designed population is depended on the crop products produced from its land, selecting an As tolerant crop to cultivate in As contaminated land could contribute directly to the livelihood of people living in As contaminated environment. Furthermore, it is also essential for the financial success of this heavily contaminated situation. Ideally, the selected crop(s) will able to survive in high As-polluted environments, be economically viable and yield a product that has either a non-food industrial application where biologically dangerous levels of As are not problematic, or a food-product with no, or safe, levels of As contamination.

1.4 Using *Brassica juncea* as a sustainable agronomic solution

B. juncea was chosen for this study as a potential As tolerant crop for the Bengal Delta due to its bioremediation potential and the economic value of its oil. *B. juncea*, which can also be called Indian mustard, is an oil type *Brassica* with reported As tolerance (Karimi et al. 2009; Pickering et al. 2000). *B. juncea* is a heavy metal accumulator with high biomass, making it a good candidate for application in phytoremediation strategies (Heiss et al. 2003). Additionally, *B. juncea* is grown widely in Bangladesh as a major source of edible oil (Ali et al. 2008). *B. juncea* belongs to the Brassicaceae plant family, a genus which displays vegetable and oilseed varieties (Hemingway 1976). The species is highly productive with a maximum average oil yield of nearly 1 tonne/hectare (Ali and Shah 1982). Mustard oil is also major edible oil in India and has medicinal importance as well (Ahamd et al. 2015).

Numerous studies have established the tolerance of *B. juncea* to As. Importantly, this tolerance is under genetic control, probably involving multiple genes each providing a small effect (Rahman et al. 2012). The BjPCS1 gene, encoding a phytochelatin synthase-related protein from *B. juncea*, has been reported as being involved in As detoxification. Overexpression of this gene could increase the level of thiol-rich peptides leading to an increase in As tolerance (Heiss et al. 2003).

The complex hybrid genome of *B. juncea* has caused challenges in studies and complicated almost all genetic analysis of the trait, and in particular, the identification of potentially As tolerant genes. *B. juncea* is a natural amphidiploid (Figure 1.4, AABB genome, 2n=36) derived by hybridization between *Brassica rapa* (AA genome, 2n=20) and *Brassica nigra* (BB genome, 2n=16) (Prakash et al. 2009).

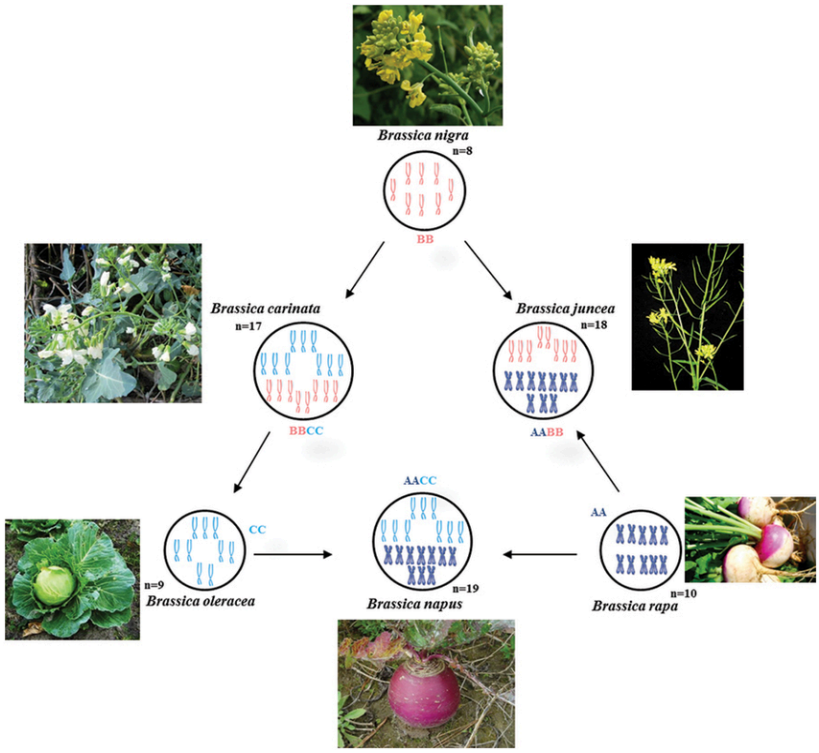

Figure 1.4. Triangle of U diagram illustrating the genetic relationships between six species of the genus *Brassica* (Kumar et al. 2015).

To address this issue, one alternative can be the use the model plant *Arabidopsis thaliana* (*A. thaliana*) to study the genetic and molecular mechanisms behind plant As tolerance, because it has a much simpler and smaller genome. Since *A. thaliana* is also a member of the Brassicaceae family, therefore, the knowledge gained could potentially benefit from substantial colinearity with related species such as *Brassica* spp., since current genomic information can provide detailed information of genetic synteny among theses species (Kowalski et al. 1994a; Kaul et al. 2000; Silady et al. 2011). *A. thaliana* is a wild species widely distributed in diverse environments and contains resources to allow efficient quantitative analyses aimed to identify the genetic and molecular based adaptation (Silady et al. 2011). Overall, *A. thaliana* may provide a viable route not only to increase the understanding of As tolerance in plants, but also to identify useful genes in improving As tolerance and bioremediation efficiency of *B. juncea*.

Other advantages of using *A. thaliana* are discussed in the following sections. The

understanding of the molecular metabolisms and genetic basis of As tolerance in *A. thaliana* can provide information to search for candidate genes and molecular markers, therefore contributing to plant breeding. Aside from its genome attributes, unlike *B. juncea, A. thaliana* is self-compatible and can be inbred easily, thus simplifying genetic experiments. Another advantage over *B. juncea* is the ease of making and controlling the genetic crossovers.

1.5 The molecular metabolism of arsenic in plants

Plants have developed a multiplex network of highly efficient homeostatic metabolisms that serve to control the uptake, accumulation, trafficking, and detoxification of As (Figure 1.5) (Yang and Chu 2011). After sensing As molecules, a series of metabolic steps occur in plant roots. These include root cell As uptake, root cell As detoxification, long-distance xylem transfer and shoot cell sequestration of As. Different As forms can be processed by plants via different metabolic pathways.

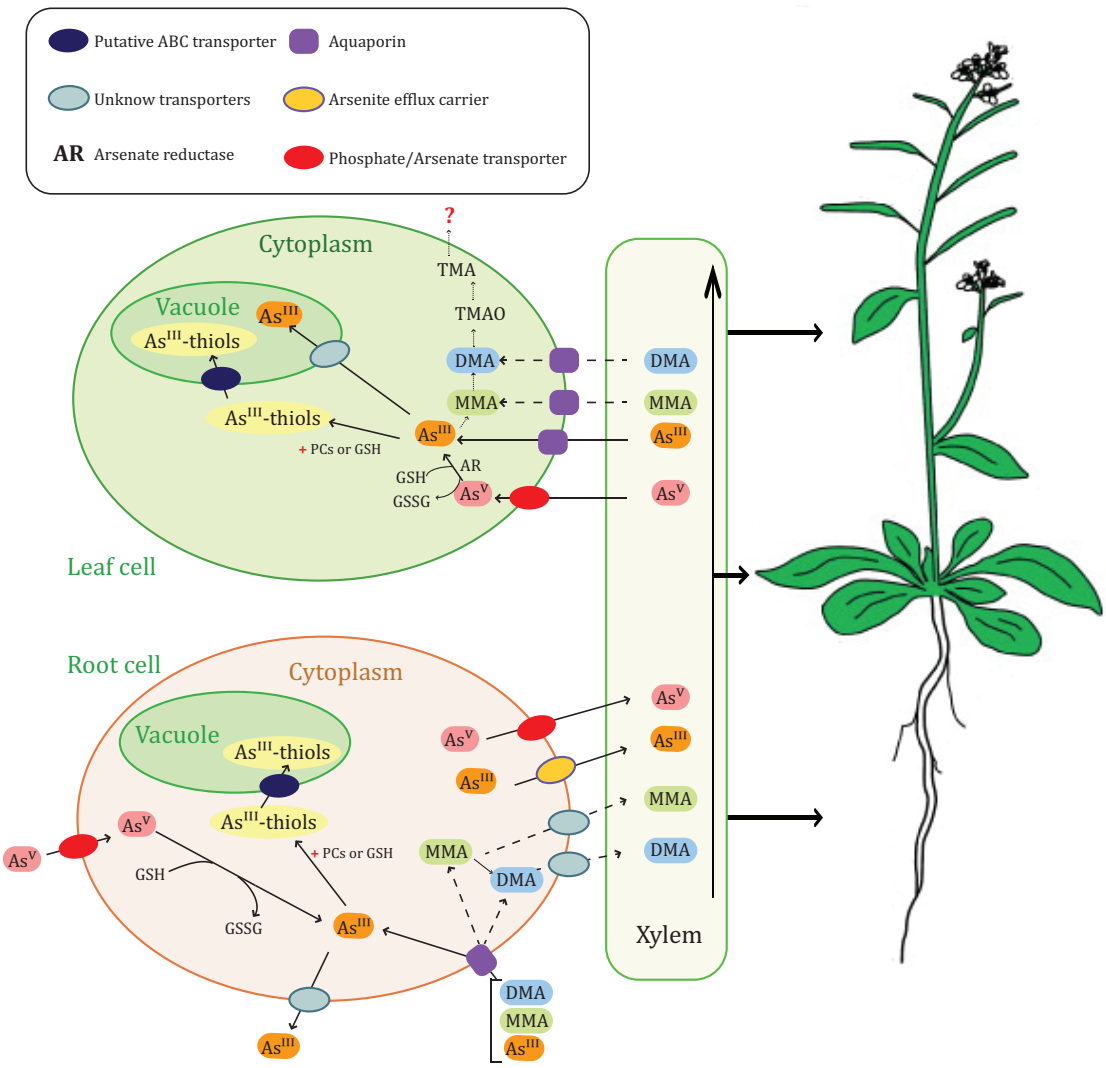

Figure 1.5. An illustration of As transfer and detoxification pathways *in planta*. (Figure adapted from Li et al. 2009; Yang and Chu 2011 and Zhao et al. 2010). Accumulation of As in the aboveground tissues of plants is determined by several factors, including the bioavailability of As in the soil, uptake from the soil solution by roots and radial transport within the root to the vascular system, translocation from the root, and storage in the above ground tissues. Arsenic species explained in the figure include arsenite (As^{III}), arsenate (As^{V}), monomethylarsonic acid (MMA) and dimethylarsinic acid (DMA).

1.5.1 Arsenate toxicity and uptake in plant

More progress has been made in the understanding of arsenate (As(V)) uptake in

plants than in other chemical forms of As. Arsenate is the leading species in aerobic soils, yet its concentration in soil solution is usually low. Less than 50nM arsenate in uncontaminated and moderately contaminated soils and 2.3 μM in a highly contaminated soil (Fitz and Wenzel 2002). This is attributed to the high affinity of arsenate for iron oxides/hydroxides in soil (Li et al. 2009). Studies in structural chemistry, plant physiology and electrophysiology have confirmed that arsenate is a structural analog of phosphate (Figure 1.6).

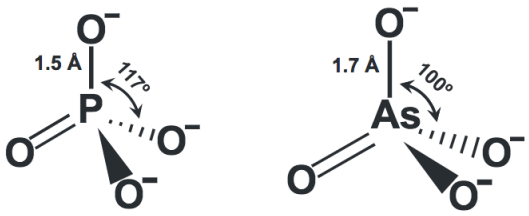

Figure 1.6. Phosphate and arsenate are structurally similar**.** Chemical structure of phosphate (left) and arsenate (right) (Rosen et al. 2011).

Because of the similarity between arsenate and phosphate, the former interferes with the biochemical pathways of the latter, such as phosphorylation and ATP synthesis (Zhao et al. 2010). The phytotoxic responses of arsenate on the plants cause the increases of As accumulation, lipid peroxidation, and *in vivo* H_2O_2 contents in roots. Lipid peroxidation is the response of plants to arsenate to induce hydroxyl radical (HO.), superoxide radical (O_2.-) or hydrogen peroxide (H_2O_2) (Ahsan et al. 2008; Ercal and Gurer-Orhan 2001).

Arsenate shares the phosphate transport pathway for uptake in higher plants, but with a lower affinity than phosphate (e.g. Ullrich-Eberius et al. 1989; Meharg et al. 1994; Li et al. 2009). When the soil contains arsenate, plants will absorb arsenate instead of phosphate by “mistake”. Extensive research confirmed that phosphate transporters play an important role in arsenate resistant in plant. Shin et al. (2004) identified two phosphate transporters, Pht1;1 and Pht1;4, both of which had a significant function in phosphate acquisition. By testing the arsenate tolerance of an *A. thaliana* double

mutant *pht1;1/4*, Bucher (2007) confirmed phosphate transporters could mediate arsenate uptake. González et al. (2005) also found an *A. thaliana* mutant *phf1*, which has inoperative phosphate transporter traffic facilitator 1 (PHF1), and fails to transfer the Pht1;1 protein from the endoplasmic reticulum to the plasma membrane. Thus, plants with higher arsenate tolerance in plants is probably due to fewer membrane transporters, lower ratio of affinity between phosphate and As, or less efficient transporters, thus less accumulation of As. However, arsenate tolerance could also be due to more efficient detoxification mechanisms. The study of an *A. thaliana* mutant *pht1;1-3* revealed an intriguingly dual phenotype. When knocking out a gene encoding for phosphate transporter PHT1;1, reduced arsenate uptake in the short-term and more As accumulation in long-term was observed in the mutant (Catarecha et al. 2007).

1.5.2 Arsenite toxicity and uptake in plants

The trivalent arsenicals, arsenite and methylated form of As have more toxic properties than the pentavalent Asals (Isayenkov and Maathuis 2008). Arsenite can bind to vicinal sulfhydryl groups of proteins to change their structures and to affect catalytic functions (Hughes 2002). Arsenite stress increases the expression of a number of enzymes that are involved in the antioxidant responses, which has been confirmed through examining the root proteome change under arsenite stress in maize (*Zea mays*) (Requejo and Tena 2005). The first enzyme involved in the As detoxification process is superoxide dismutase (SOD), which changes superoxide radicals (O_2) into hydroxide peroxide (H_2O_2). Then catalase (CAT) and peroxidases (POD) catalyze the scavenging of H_2O_2. Ascorbate peroxidases (APX) act together with glutathione reductase (GR) and participate in the ascorbate-glutathione cycle, which additionally promotes hydrogen peroxide decomposition (Requejo and Tena 2005; Leão et al. 2014).

Arsenite is the dominant As species in anaerobic environments such as submerged

soils (Li et al. 2009). Until relatively recently, little progress has been made in the area of arsenite study. Unlike arsenate, which enters into root cells via competing for transporter with phosphate, arsenite employs plant aquaporin channels to transfer itself from the external medium into roots, and this pathway is impeded by glycerol and antimonite (Li et al. 2009). Research in the yeast (*Saccharomyces cerevisiae*) suggested undissociated arsenite was transported across the plasma membrane via a glycerol-transporting channel; then, a competitive study was performed among arsenite, glycerol and antimonite in rice and found glycerol and antimonite competed with arsenite for transportation in a dose-dependent manner (Meharg and Jardine 2003). The same phenomenon was then found in wheat and pea root cells, indicating that arsenite and glycerol uptake share the same route *via* aquaporins (Chaumont and Tyerman 2014). Since then, understanding of the arsenite uptake pathway has extended further. A subfamily of the plant aquaporin family called nodulin26-like intrinsic proteins (NIPs) was discovered to play a significant role in arsenite transference, but so far, the aquaporin channel is the only pathway recognised for arsenite uptake (Wang et al. 2015).

Recent studies in rice established that arsenite shares the silicon transport pathway for entry into rice root cells (Li et al. 2009). When arsenite converts to arsenous acid, it has a similar molecular size, identical tetrahedral forms, as well as the comparable high pKa to silicic acid (Ma et al. 2006). Rice has very high affinity to silicon (Zhao et al. 2010). Arsenite appears to enter rice roots through a silicon efficient transporter, OsNIP2;1 (also named Lsi1; Ma et al. 2006). Therefore, rice accumulates a relatively high concentration of As leading to a food safety concern. This discovery is highly relevant to the public concerns of As pollution in Bangladesh, because rice is largely grown and consumed in Bangladesh but should arguably be replaced by an alternative low As affinity crop.

1.5.3 Detoxification of arsenic in plants

Almost all organisms have evolved mechanisms to detoxify As. Although some plants can restrict arsenate uptake through suppression of high-affinity phosphate transporters, most plants have a number of transporters and enzymes specific for conferring As tolerance, or transform As into a less toxic As species (Rosen et al. 2011). Another possible constitutive mode of detoxication is the efflux of As to the external medium, but these studies are mostly conducted in bacteria, and there is no clear evidence in plants so far (Yang et al. 2012). Two well-studied metabolic modes of detoxification are As complexation with thiol compounds into non-toxic form (chelation) and vacuolar sequestration *in planta.* The detail of this pathway has also been presented in figure 1.5.

Chelation

Chelation is considered a typical heavy metal detoxification process in plants and other organisms. Toxic elements are bound with a peptide ligand, such as metallothioneins (MTs), small gene-encoded cysteine-rich polypeptides, and in some cases, then subsequent compartmentalization of the ligand-metal complex to convert into non-toxic elements and then are removed from the cytosol (Cobbett 2000).

Phytochelatins and metallothioneins are the two different classes of ligands. Phytochelatins are enzymatically synthesised cysteine-rich and glutathione-derived peptides. They have a general structure (γ-Glu-Cys)n-Gly (Cobbett 2000; Mendoza-Cózatl et al. 2011). Phytochelatins and glutathione (GSH) play an important role in heavy metal detoxification (Zhao et al. 2010). Metallothioneins are gene-encoded polypeptides and play a more crucial role in human As tolerance (Lee et al. 2010), while phytochelatins are more likely employed in plants. Blocking phytochelatins synthesis results high level of both arsenate and arsenite sensitivity. Sensitivity of *A. thaliana* PC-deficient mutant was increased 10–20 times compared to wild type (Schat et al. 2002). In *A. thaliana*, two phytochelatin encoding genes have been reported. Vatamaniuk et al. (1999) used the expression of *A. thaliana* cDNA

libraries in *S. cerevisiae* and identified gene AtPCS1. Another PC synthase gene, AtPCS2, was identified in *A. thaliana,* as it is highly similar to AtPCS1 (Ha et al. 1999). But in most tissues, AtPCS2 is expressed at a relatively low level compared with AtPCS1, so the physiological function remains to be determined (Cobbett and Goldsbrough 2002). When plants are exposed to arsenate, genes and enzymes involved in glutathione synthesis, metabolism and transportation are typically up-regulated (Ahsan et al. 2008). Enhancing the synthesis of thiol also increases As tolerance. It also has been confirmed that large amounts of arsenite in plant tissues is complexed with thiol-rich peptides into a nontoxic form. The detail of this pathway has also been presented in figure 1.5.

The detoxification of arsenate is through chemical reduction to arsenite, followed by arsenite detoxification (Dhankher et al. 2002; Pickering et al. 2006; Raab et al. 2005). Dhankher et al. (2002) found 96% – 100% of the As in *A. thaliana* was reduced to arsenite in wild type shoots of three-week-old seedlings, when exposed to arsenate. Employing extended x-ray absorption fine structure analysis, Pickering et al. (2000) found that the reduction of arsenate (AsV) to As(III) in the Indian mustard roots is coordinated by three sulfur ligands, known as the AsIII-tris-glutathione complex.

The phytochelatins synthase is strongly activated by As stress or stress from exposure to highly toxic ions of cadmium, lead, mercury, and other metals (Cobbett 2000; Schmöger et al. 2000). Thus, phytochelatin increases when the cell suffers high concentrations of metal ions (Vatamaniuk et al. 2000). However, the efficiency of the enzyme binding to metal is not correlated to its catalytic activation. Although AtPCS1 has a much higher affinity and capacity for cadmium than to other metal ions, like copper, it can equally be activated by these ions (Vatamaniuk et al. 2000). Additional studies are required to establish the precise influence of this pathway on PC function.

Vacuolar sequestration

All excess heavy metals, apart from iron, are stored in the vacuole (Martinoia et al. 2012). Arsenic and cadmium, which do not serve in plants as micronutrients, are stored in vacuoles due to the unselective character of some transporters (Martinoia et al. 2012). Investigation of As distribution is challenged at the cellular and subcellular level in non-hyperaccumulator plants because current elemental mapping techniques are not sensitive enough for the comparatively low concentrations of As in plant tissues, or it is hard to provide electron microscopy images with sufficiently high resolution. Another big challenge is to quantify the localization of different As species (Li et al. 2009).

Some evidence suggests trace arsenite is taken up into vesicles, but only 20% of them will transform into As(III)-$(GS)_3$ (Li et al. 2009). By analogy, some studies analyze As and cadmium together. Li et al. (1997) first discovered a vacuolar member transporter in yeast, a cadmium factor 1 (YCF1) conferred resistance to cadmium and As. YCF1 is an ATP-binding cassette (ABC) transporter and can regulate the transportation of glutathione-S-conjugated metals and metalloids, such as Cd-$(GS)_2$ (Li et al. 1997) and As(III)-$(GS)_3$ (Ghosh et al. 1999). Because of the likely large concentration gradient from the cytoplasm to the vacuole, transport of arsenite across the tonoplast probably involves an energy-dependent active metabolism (Li et al. 2009).

Song et al. (2010) reported the identification of two ABC transporters, AtABCC1 and AtABCC2, which are required for As detoxification. Although the lose-of-function single mutant of AtABCC1 or AtABCC2, did not show a significant difference compared with the wild type in the presence of As [V] or disodium methanearsonate (DSMA, from an As-containing herbicides), the double mutant were extremely sensitive to As(V) and DSMA, when the concentrations of As and DSMA were both above 30μM. Furthermore, the transport activities for As(III)–PC_2 and apoPC_2 into

intact vacuoles isolated from the double mutants were dramatically reduced by 85% and 95%, respectively, compared with the wild type. Thus, these data indicated that these two ABC transporters are by far the most important PC and PC-conjugate transporters in *A. thaliana*. However, transgenic plants with overexpressed AtABCC1 gene have increased transcript levels of maximally 20–40%, but no increased As tolerance. The AtPCS1 and AtABCC1 co-overexpressing lines exhibited significantly consistent increase in As tolerance and shoot fresh weight. To conclude, overexpression of vacuolar member transporter AtABCC1 increased As tolerance only when co-expressed with PCS.

To date, As detoxification in plants has mostly been investigated through biochemistry approaches, with little focus on understanding As specific detoxification mechanisms at the genetic level. Many reported molecules in As transfer and detoxification are more likely to deal with general heavy metal tolerance. Thus finding As tolerance-specific genes, as well as improving the understanding of the genetic basis of As is necessary. One approach is to search for genes, and their functional nucleotide polymorphism as underlying allelic variation, that account for plant interspecific natural variation for quantitative and qualitative traits.

1.6 Understanding the genetic basis of arsenic tolerance through genetic mapping

Genetic mapping is the common name for all approaches used to identify the locus of a gene and the distances between genes (Oraguzie et al. 2007). It is also known as linkage mapping or meiotic mapping, which means placing the markers in chromosal liner order, indicating the relative genetic distance between them and assaying them to their linkage groups by recombination values from all pairwise combination of the markers.

The precondition for QTL mapping is that the trait of interest needs to be a quantitative trait. Many agriculturally important characters such as yield, quality, and some forms of disease resistance are controlled by multiple genes or quantitative loci. In As tolerance studies, rice (*Oryza sativa*) and maize (*Zea mays*) were used to analyze the As accumulation trait. In the study of maize, the As content in different tissues, including kernels, axes, stems, bracts and leaves, were examined. After phenotyping a mapping population of 194 F8 generation recombinant inbred lines (RILs), twenty-eight QTLs were found with an association to As accumulation in different tissues (Fu et al. 2016). A double haploid population was used in another study of rice and four QTLs involved in As accumulation in different plant developmental stages were found (Zhang et al. 2008). These results supported the feasibility of studying the quantitative traits of As tolerance.

The process of QTL mapping involves four major steps; 1) development of a mapping population, 2) generating a full linkage map, 3) phenotyping of the mapping population, and 4) QTL detection using statistical tools.

1.6.1 Mapping population selection

Genetic mapping begins the creation of a new, or selection of an existing, mapping population. Plant mapping populations are usually created from F1 lines, which are derived from two parents with different phenotypes for a target trait. For example, one *A. thaliana* genotype with high root elongation ratio and another with low root elongation rate under the same As stress were used to indicate a suitable mapping population (Sánchez-Bermejo et al. 2014). Randomly crossing F1 individuals or by allowing self-pollination can generate a segregating F2 population. These F1 individuals are heterozygous, and homozygous lines can be then produced by using single seed descent, which means one seed harvested from each F2 individual is used for growing F3 plant and so on until the F8 or F9 generations (Grant and Shoemaker 2009).

The classic mapping population includes two groups. One is bi-parental mapping population (linkage mapping). Recombinant inbred line (RIL) population and the doubled haploid line (DH) populations belong to this group. They are also considered as the primary mapping populations. The secondary population is generally developed from primary population, such as a near-isogenic line (NIL) population and a chromosome segment substitution lines (CSSL) population.

Another main group is the association mapping population. As a model plant, *A. thaliana* accessions have been widely used to generate RIL populations for different research purposes. Up to 39 RIL populations are available via TAIR (the *A. thaliana* information resource web). These resources can be particularly useful once these lines have been genotyped. They can then be phenotyped multiple times, making it possible (as well as extremely cost effective) to study many different traits in many different environments (Fournier-Level et al. 2010). Generally, there are less than 200 lines in a bi-parental mapping population, but the population as the Multi-parent Advanced Generation Inter-Cross (MAGIC) Lines, contains 527 RILs with the corresponding sequenced genomic information (Kover et al. 2009).

Since there are many available mapping populations, the usual method to minimise mapping time is to choose one population, which has a phenotypic variation of a segregated trait derived from their parental line, rather than to cross suitable chosen parents to develop new progeny families. In the study of As tolerance, if there is a mapping population generated from two parents with different As tolerance level, it is likely this population would have As tolerance segregation (Rieseberg et al. 1999).

1.6.2 Accuracy of QTL mapping

The success of a QTL mapping depends on its mapping resolution. By exploration of genomic data on corresponding regions around the peak QTL together with information from within the region delimited by flanking markers, researchers can

identify candidate loci of interest (Tecle et al. 2010). The accuracy of a QTL mapping is related to the distance between marker and a peak QTL. Moreover, it is dependent on the type of mapping population, the number of individuals scored, and the quality of the data (Price 2006). Generally, initial positioning of a QTL could cover a region in excess of 20cM (Kearsey and Farquhar 1998). However, a QTL mapping study in *A. thaliana* is relatively less challenging than in other plant species and in some cases, targeted plant QTLs might cover regions as small as 2cM or less (Price 2006). The strategy for fine mapping follow-up is to genotype a large number of SNPs (more than 1,000) across the peak of the admixture association, at a resolution of one every few thousand base pairs. Using a secondary mapping population to perform the fine mapping is another option, but is time-consuming, expensive and practically difficult if the QTL effect is small, because the small genetic effect limits the ability to assess phenotypic difference accurately between allelic variants in the fine mapping population. Most small QTLs will not be tractable using a secondary mapping population, but the original QTL mapping is probably more accurate than was previously supposed (Price 2006). Another complex phenomenon that can interfere with the mapping accuracy is the occurrence of two or more closely linked QTL, suggesting that some traits might have a more complex polygenic architecture than that determined by QTL analysis (Kroymann and Mitchell-Olds 2005).

1.6.3 Candidate gene selection

The construction of detailed genetic maps with high levels of genome coverage is the first step to localise genes or QTL that are associated with economically important traits. Once QTLs have been identified, the next step is to determine the identity of genes that contribute to the phenotype. One of the most promising ways to do this is by positional cloning, where the QTL is linked to the physical sequence of the genome via the sequence of large insert clones (Salvi and Tuberosa 2005). *A. thaliana* has been sequenced and gene order is already known, so there is no need to generate large insert clones but move to gene discovery and marker-assisted selection (Collard

et al. 2005; Grant and Shoemaker 2001). Gene scanning can also be benefited from the plants with the known molecular metabolism in the presence of As or other heavy metals. The potential genes could be the polymorphisms located in the coding region and are more likely to be related to As transfer or detoxification process. Once the gene or genes have been identified, some typical methods can be used to alter the gene expression (Alonso-Blanco et al. 2005). However, the stress of heavy metals plays an mild role in natural selection, therefore, traits are more likely to be caused by small QTLs, where its contribution to overall variation is in the order of 25% or less (Price 2006).

1.7 Genome-wide association studies

Studies of genome-wide association (GWA) in human genetics has contributed significantly to the understanding of the genetic basis of common human diseases, and GWA has been a useful tool ever since (Platt et al. 2010). With the increased resources of plant genotyping and sequencing data and technology, GWAS analysis has become a distinct general approach for studying the genetics of natural variation and traits of agricultural importance. Integrating bi-parental mapping and GWAS can be used as a powerful complementary approach to classical bi-parental mapping to dissect complex traits (Korte et al. 2012). However, QTL mapping studies are sometimes restricted in allelic diversity and have the limited genomic resolution. The genomic resolution was affected by the size of the population. GWAS analysis overcomes this limitation, as it uses unlimited sample size, as long as the genotyping information is available (Korte and Farlow 2013). In a maize rough dwarf disease study, 236 maize inbred lines were firstly used. The target region was then narrowed down into 41,101 single-nucleotide polymorphism (SNP) markers and a GWAs analysis found a total of 73 SNPs were associated with rough dwarf disease resistance (Liu et al. 2014). In another study in soya bean, 141 inbred lines were used first and 778 SNPs were derived from QTL analysis, then 304 soya bean genotypes were used for GWAs analysis and eight genomic loci associated with more complex

traits (maturity, plant height, seed weight, seed oil and protein expression) were discovered (Sonah et al. 2015).

1.8 Arsenite tolerance trait selection

I believe the biggest challenge in the study of genetic variations and phenotypic diversity of As tolerance is to find a physical parameter to represent As tolerance trait, which can show a consistent and statistically significant phenotypic difference that is caused by As. In many studies of abiotic stress, the relative growth rate (with and without stress) was employed (Hunt et al. 2003). When measuring "growth", researchers need to be careful with the role of different processes at different times, because the growth rate of a plant is the result of the effect from both the environment in which it grows and its genetic background (Fourcaud et al. 2008). In previous research, plant growth analysis used many simple primary data from measuring weight, areas, volumes and content of plant component to analyze process within and involving the whole plant (Evans 1972). A suitable parameter to compare the tolerance level of an individual accession and discover the underlying linkage between genotype and phenotype is the key to a successful mapping. Generally, under As stress, plants are much smaller in size and tend to have extremely short roots. Consequently, root elongation rate is the most frequently used phenotype. Plant phenotypes, such as root length (Beemster et al. 2002; Mouchel et al. 2004), leaves size and numbers (Pérez-Pérez et al. 2002; Horiguchi et al. 2006), seeds size and numbers (Alonso-Blanco et al. 1999; Doganlar et al. 2000), or carpels and fruits (Tanksley 2004), have natural variations which are associated with the variation in cell size and number. Detailed and dynamic analysis of growth parameters such as root elongation in *A. thaliana* (Beemster et al. 2002) shows that natural genetic variation affects cell division and cell growth while genetic differences in cell growth affect cell cycle duration and the number of proliferating cells. The change of root length, shoot weight as well as shoot pigments should all be examined to find a correlation with plant arsenite tolerance ability.

1.9 Research schemes

The first task in this project was to test the arsenite tolerance and absorption in the oil crop *B. juncea.* Although *B. juncea* does not have a super As accumulation ability, it is an important oilseed crop in Bangladesh, and therefore it was necessary to establish whether its seeds can accumulate As or not, and if so, how much As is present in the seed-extracted oil. The results described in Chapter 3 clearly demonstrated that As accumulated to levels above 10ppb in the seed oil of *B. juncea* plants grown in the presence of environmentally-relevant concentrations of As. The data also established that the arsenite tolerance and absorption ability varied among the different *B. juncea* ecotypes.

This outcome guided the direction of the project toward discovering 1) how to improve the As tolerance of crop plants and crucially, 2) to identify potential ways to reduce or prevent the accumulation of As above safe levels in the oil of plants growing on contaminated land.

To study the genetic basis of arsenite tolerance variation, the arsenite tolerance of multiple *A. thaliana* land races (genotypes) was analyzed and confirmed that arsenite tolerance variation can be measured as a plant shoot weight change and this change is under genetic control. The arsenite tolerance genotype and sensitive genotype were cataloged for genetic mapping analysis.

The genetic analysis of arsenite tolerance was conducted in Chapter 5, which was employed genetic studies and crop improvement methods to increase the understanding of plant arsenite tolerance. The loci directly associated with plant arsenite tolerance were identified in this chapter. Information of quantitative trait loci obtained from genetic analysis could be applied in marker-assisted selection to improve the arsenite tolerance of crop plant, such as *B. juncea*.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Plant Materials

2.1.1.1 *Brassica juncea*

Seed stocks of 14 Bangladesh local ecotypes including Bar-11, BJ-17, Daulat, Jun-30, Jun-57, Jun-165, Jun-895, Jun-1495, GP-Bhaibob, GP-Jun-401, GP-Jun-517, GP-Jun-1197, GP-Varioa and Rai, were kindly provided by Dr. Nazmul Haq (University of Southampton, UK).

2.1.1.2 *Arabidopsis thaliana*

Seed stocks of 19 ecotypes (Table 2.1) were used as parental lines to generate a Multi-parent Advanced Generation Inter-Cross (MAGIC) population (Kover et al. 2009). These were kindly provided by Dr. Paula Kover (University of Bath, UK) (Table 2.1).

Table 2.1 Information of 19 parental lines of MAGIC population.

Accession	Origin	AIMS Stock Centre
Bur-0	Ireland	CS6643
Can-0	Canary Isles	CS6660
Ct-1	Italy	CS6674
Edi-0	Scotland	CS6688
Hi-0	Netherlands	CS6736
Kn-0	Lithuania	CS6762
Ler-0	Poland, formerly Germany	CS20
Mt-0	Libya	CS1380
No-0	Germany	CS6805
Oy-0	Norway	CS6824
Po-0	Germany	CS6839
Rsch-4	Russia	CS6850
Sf-2	Spain	CS6857
Tsu-0	Japan	CS6874
Wil-2	Russia	CS6889
Ws-0	Russia	CS6891
Wu-0	Germany	CS6897
Zu-0	Germany	CS6902

Seed stocks of other 17 ecotypes (Table 2.2) were obtained from the Nottingham *Arabidopsis* stock center, and 4 others were generated in our laboratory by Prof. Rod Scott (University of Bath, UK), which were originally ordered from Nottingham *Arabidopsis* stock center (Table 2.2, labeled with *).

Table 2.2 Information of 21 *A. thaliana* used in this research study.

Accession	Origin	AIMS Stock Centre
Bla-1 *	Spain	CS970
Bolin-1	Romania	CS76373
Col-3	U.S.A	CS908
Col-5 (gl1)		CS1644
Cvi-0 *	Cape Verde Islands	CS1096
C24 *	Portugal	CS906
Fei-0	Portugal	CS28250
Gr-1	Austria	CS1199
Kas-1	Kashmir	CS903
Kas-2	Kashmir	CS76150
Mer-6	Spain	CS76414
Mrk-0	Germany	CS28498
Nd-1	Germany	CS1636
Ob-0	Germany	CS1418
Ri-0	Canada	CS28686
Rld-1 *	Netherlands	CS913
Rou-0	France	CS28692
Rovero-1	Italy	CS76351
Rsch-0	Russia	CS1490
Sf-0	Spain	CS1510
St-0	Stockholm	CS6863
Valsi-1	Italy	CS76425

A set of 154 recombinant inbred lines (RILs) (AIMS Stock Centre: CS39695) was generated from the cross between the accessions Ler-0 as the seed parent and Kas-2 as the pollen parent (El-Lithy et al. 2006), and was obtained from the Nottingham *Arabidopsis* stock center.

The At2g21045 T-DNA mutant (GK-868F11, Col-0 background) (Sánchez-Bermejo et al. 2014) was obtained from the Nottingham *Arabidopsis* stock center.

2.1.2 General laboratory reagents

All regents for preparing Hoagland solution were supplied by Sigma-Aldrich (Gilligham, UK).

Ethanol was from Fisher Scientific (Loughborough, UK). Double-distilled water ($_{dd}H_2O$) and ultra-purified water (MilliQ water) was produced in the lab from an Elix system and a Milli-Q Integral syatem (Millipore (UK) Ltd, Watford, UK), respectively.

2.1.3 Hydroponic growth media

The hydroponically cultured *B. juncea* were grown in a ready-to-use GH Floraseries solution (Greens Hydroponics Ltd, UK), with $_{dd}H_2O$, with the pH adjusted to 7.2 with 1M NaOH.

Figure 2.1. GH Floraseries solution.

The hydroponically cultured plants of *A. thaliana* were grown in the Hoagland solution (Hoagland and Arnon 1950) with the pH adjusted to 5.7 with 1M HCl or 1M NaOH. The recipe of Hoagland's complete nutrient solution is presented in the supplementary data 1.

2.1.4 Plant seedling growth medium

Plant seeding growth medium for both *A. thaliana* and was prepared with 1/2 strength Murashige and Skoog medium (MS) with vitamins (Murashige and Skoog, 1962) (Melford Laboratories Ltd, Suffolk, UK), 1% (w/v) sucrose and 0.8% (w/v) agar supplement. The pH of the medium was adjusted to 5.7 with 1M HCl or 1M NaOH.

2.1.5 Designing and synthesis of primers for polymerase chain reaction

Primer3 web (version 4.0.0) was used to check all primers and to determine melting temperatures. After a primer was designed, a megaBLAST (NCBI) search for highly related nucleotide sequences across the entire *A. thaliana* genome was carried out to ensure target specificity. Primers were ordered from Primerdesign (PrimerDesign Ltd, Southampton, UK). Primer information is given in Table 2.3.

Table 2.3 Primer sequences used for genotyping and expression analysis

Primer	Sequence (5'-3')	Tm (^{o}C)
AtPCS1 RT-PCR For	CTTTGGGAAGCCATGGACAG	54
AtPCS1 RT-PCR Rev	TGTTCATTCCCATCACGCAC	54
At5g04885 For (P1F)	ACGGGTCTGGTCTTGAAACT	54
At5g04885 Rev (P1R)	AGTCCTCGTATTGAACCCCG	54
At5g05390 For (P2F)	AAAATTCCCTAAAAGCCGGTG	55
At5g05390 Rev (P2R)	GGTAGATCGTGAGGAGGAGC	55
At5g05780 For (P3F)	ACATTCTCAACATGTGGGC	54
At5g05780 Rev (P3R)	CGACGGAGCTTGAGAGTTTG	54
At5g06060 For (P4F)	AAGAGAGGAGCGACAAACCC	54
At5g06060 Rev (P4R)	AACGTTGTCTAAAAGAGGAAACC	54
At5g06410 For (P5F)	CCCAGAGCTACAGATTTTAAGGT	53
At5g06410 Rev (P5R)	TGGCAAAAGAAGCTACATCCTG	53
At5g06805 For (P6F)	TGTGATGGGGTCGATCTCTG	54
At5g06805 Rev (P6R)	TGCAACGTCTTTACAACCTGA	54
At2g21045 For	GGGCTAGTTGTTTCTCCTTGC	53
At2g21045 Rev	AAGCACGACTTCCTCTTCCTC	53
GAPC For	CACTTGAAGGGTGGTGCCAAG	55
GAPC Rev	CCTGTTGTCGCCAACGAAGTC	55

2.1.6 SNP marker for fine mapping

The SNP and indel differences in the genome of a target region on chromosome 5 between Ler-0 ecotype and Kas-2 ecotype genome were used to design the molecular maker for fine mapping.

a. Six pairs of indel difference markers were designed based on the different insertion and deletion variations between Ler-0 ecotype and Kas-2 ecotype. Each primer was targeting a fragment of 800-1000bps, which contained a difference of 20-40 bps.

b. To design the SNP makers, four or five options of sequences were required for probe synthesis. Each sequence had one or two SNP differences, and a length between 200bps to 500bps. After the sequence was selected, a nucleotide BLAST (NCBI) search for verifying sequence uniqueness across the entire *A. thaliana* genome was carried out to ensure target specificity. RepeatMasker (webtool) was applied to screen selected sequences to avoid any interspersed repeats and low complexity DNA sequences. SNP genotyping probes were synthesised by BGI (Hong Kong, Chian) using sequences containing SNP differences of 30 sites with at least 4 options of each site (Liu et al. 2012).

2.1.7 Statistical analysis

Statistical analyzes were performed using SPSS 22 software. Comparisons for normally distributed data with equal variance were carried out using t-test, one-way analysis of variance (one-way ANOVA) and two-way analysis of variance (two-way ANOVA) followed by Tukey's multiple comparisons (family error rate = 5). Data not following a normal distribution was analyzed using the Mann-Whitney U-test with *P-value* adjusted for ties.

2.1.8 Image capture and processing

Photographs of whole plants and plant seedlings were obtained using a V-100 digital camera (Canon). Mature seeds were photographed under a SMZ1500 dissecting microscope (Nikon) using a Digital Sight DS-U1colour camera (Nikon). Images were processed using ImageJ-1.47V.

2.2 Methods

2.2.1 Growth conditions for *B. juncea*

Seeds of all *B. juncea* ecotypes were bulked in a glasshouse with a day length of 16 hours at 25°C and night length of 8 hours at 18°C. The growth of *B. juncea* can be classified into four stages.

(1) All ecotypes were given a stratification treatment for 4 days in 0.15% agarose at 4°C to promote synchronous germination. 10 stratified seeds were sown into pot (6cm diameter x 6.5cm height) with Levington F2 compost pre-treated with 0.2gm/L insecticidal intercept solution (Intercept 70 WG - Scotts). Sown seeds were covered with propagator lids.

(2) After seed germination and shoot attainment of 2cm height, each seedling was transplanted into a pot (5cm length x 5cm wide x 5.5cm height) filled with Levington M2 compost, until the roots grew out through the holes under the pots.

(3) Plants were then transplanted into a new pot (8cm length x 8cm wide x 10cm height), filled with Levington F2 compost mixed with perlite.

(4) Once the roots grew out of the bottom of the pots, the plants were moved into a larger pot (18cm diameter x 15cm height) filled with Levington F2 compost and perlite with additional long lasting NPK fertilizer to provide plants with essential nutrition.

2.2.2 Growth conditions for *A. thaliana*

Seeds of all *A. thaliana* ecotypes were bulked before using in other experiments. All ecotypes were given stratification treatment for 4 days in 0.15% agarose at 4°C to promote synchronous germination. Stratified seeds were transferred into Levington F2 compost, which had pre-treated with 0.2gm/L insecticidal intercept solution

(Intercept 70 WG - Scotts). Sown seeds were covered with propagator lids for two weeks after germination. Plants were placed in a Gallenkamp growth chamber with a day length of 16 hours at 22°C in the day, and an 8-hour night length at 18°C. The humidity of the room was maintained at 60%. The tap water were given to *A. thaliana* plants when needed. In order to increase seed size in *A. thaliana*, some plants were trimmed to keep only 10 siliques.

2.2.3 Cross pollinations

For reciprocal crosses between individual *Arabidopsis* plants, the maternal plants were emasculated a day prior to anthesis and were then pollinated two days after emasculation. Pollination was done when the stigmatic papillae was fully extended. Mature pollen from the pollen parent was gently dabbed onto the mature stigmas of the seed parent. In order to avoid self-pollination and cross-pollination, the process of emasculation was gently done by selecting only 3 buds per branch, and removing 6 remaining flower buds around the emasculated buds. Both the emasculation and the pollination were carried out under a Leica MZ6 dissecting microscope. Each individual *Arabidopsis* plant for reciprocal cross was used as both seed parent and pollen parent.

2.2.4 Seedling root elongation assay of *B. juncea*

All seeds were surface sterilised in 50% (v/v) of the commercial thin bleach (containing 4.5% sodium hypochlorite) for 5 min. Seeds were then centrifuged (Eppendorf) at a low speed (2,500 xg) in a bench-top centrifuge to drain the bleach. They were then washed thoroughly with sterilised $_{dd}H_2O$ 5 times with 5 min each. Sterilised seeds were kept at 4°C for 3 days for stratification. Fifteen seeds of each accession were laid evenly on the surface of seedling growth medium in 100mm x 100mm square Petri dishes (Sterilin) and the dishes were sealed with parafilm (3M, UK) for the observation of root growth. Sodium arsenite ranging between 0ppm to

7ppm was given to seedlings as the As stress. Petri dishes were kept in a Gallenkamp growth room for 8 days. The square Petri dishes were placed vertically to allow vertical elongation of seedling roots, which was parallel to the side of the dish with each other, so the length of an individual sample was directly comparable and measurable. Images of dishes with germinated seedlings were taken and the root lengths were measured from the images.

2.2.5 Hydroponic cultivation of *B. juncea*

Clay pellets (Figure 2.2.a) and mineral wool cubes (Figure 2.2.b and c) were used as growth media. Both the mineral wool cube and the clay pellets were cleaned with $_{dd}H_2O$ with the pH adjusted to 5.7 (by using 1 M HCl or 1 M NaOH). They were then autoclaved. Seeds were germinated in Petri dishes containing a seedling growth medium in Gallenkamp growth room. Seven-day old seedlings were then transplanted into round propagation mineral wool cubes (36mm x 40mm, from Green's horticulture company) as shown in Figure 2.2.b. Once the roots of the plant grew out of the cube, the plant was transferred to a larger transplanting cube (75mm x 75mm x 75mm with a hole measuring approximately 36mm x 36mm x 40mm) as shown in Figure 2.2.c in a glasshouse.

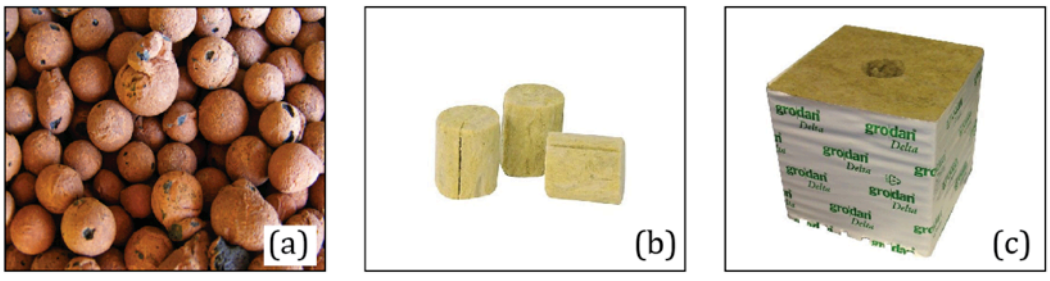

Figure 2.2. a. Clay pellets; b. propagation mineral wool cube; c. transplanting cubes.

The micro-jet pumps, which consisted of two groups of aquatrays (Green's horticulture company; Figure 2.3), were used to support the plant and to offer air and the nutrient to mineral wool cubes. The aquatrays were covered with foil to prevent

light to inhibit growth on agar.

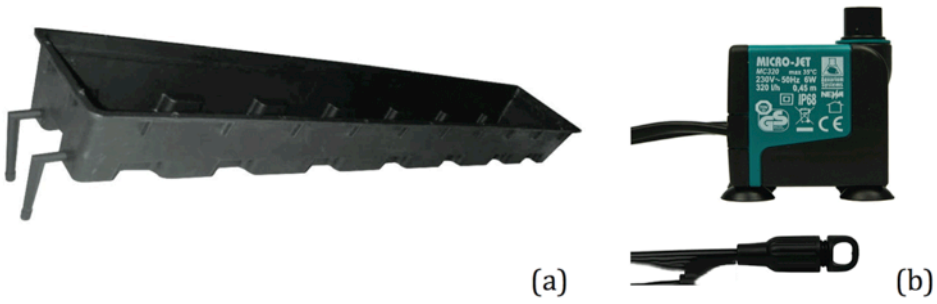

Figure 2.3. Illustrations of a. aquatray container and b. micro-jet pumps.

All equipment was located in a glasshouse with a day length of 16 hours at 25°C and a night length of 8 hours at 18°C. 1ppm of sodium arsenite was given to plants as the As stress treated group. The illustration of *B. juncea* plants growth is presented in figure 2.4 and a example of the mature plant is given in figure 2.5. Plant seeds were collected from this cultivation for further chemical analysis.

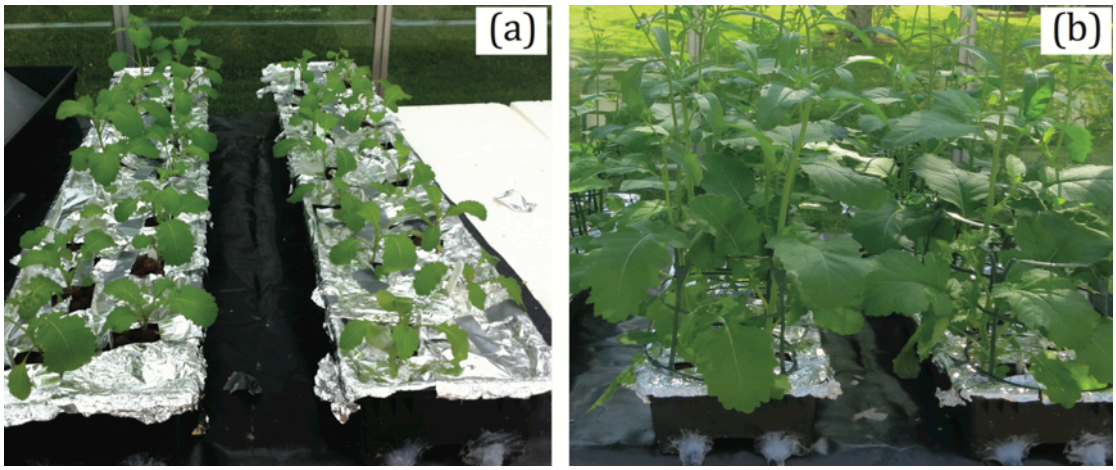

Figure 2.4. A representative image of hydroponically cultured *Brassica juncea*. a. The seedling stage. b. The bolting stage.

Figure 2.5. A representative image of a mature plant of *B. juncea* Bar-11 ecotype grown under hydroponic cultivation.

2.2.6 Oil extraction of seeds of *B. juncea*

A hexane-based oil extraction approach (Rosenthal et al. 1996; Do et al. 2014) was carried out for measuring As content of *Brassica* seed oil from hydroponically cultured plants. Seeds collected from individual plants were dried at 50^{o}C in an incubator (Unitemp) and then ground into a powder by using mortar and pestle. The weights of seed powders were measured by a Metter PC 4000 balance (Deltarange). Hexane and seed powder with a ratio of 1: 3.5 (v/w) were mixed in a 15mL sealed falcon tube and left at 75^{o}C (Unitemp incubator) overnight. On the following day, tubes were left in the fumehood with the lid open for another day to allow evaporation of the hexane. The same amount of hexane was added to the tubes to repeat the extraction step. Oil extraction samples were left in the fumehood for two to three days and the weight of samples was measured each day until the weight was constant. After centrifuging

(Eppendorf) at a maximum speed of 20,000 x g in a bench-top centrifuge, extracted oil and precipitate were separated. The upper layer of oil was transferred into a 1.5mL tube (Eppendorf).

2.2.7 Seedling root elongation assay of *A. thaliana*

All seeds were surface sterilised using 15% (v/v) of the commercial thin bleach (containing 4.5% sodium hypochlorite) for 5min followed by centrifugation (Eppendorf) at a low speed (2,500 xg) in a bench-top centrifuge. They were then washed five times for five minutes each with sterilised $_{dd}H_2O$. Twenty seeds of each ecotype were laid evenly on the surface of the seedling growth medium. Petri dishes were kept in a Gallenkamp growth room for 14 days and the length of the root was then measured under controlled conditions. The rest of the settings of this experiment was the same as described in section 2.2.4.

2.2.8 Measurements of hydroponic cultivation of *A. thaliana*

The mineral wool cubes (36mm x 40mm, from Green's horticulture company) (Figure 2.2.b) and perlite with net pot were used as growth media. The mineral wool cube was cleaned with $_{dd}H_2O$ and the pH value was adjusted to 5.7 (by using 1 M HCl or 1 M NaOH), followed by autoclaving. Perlite with net pot was soaked in Hoagland solution for 1 day before being used. The container tray was cleaned with the consortium thin bleach (containing 4.5% sodium hypochlorite). Seeds were germinated in Petri dishes containing seedling growth medium and agar in Gallenkamp growth room for 10 days. Seeds were then transplanted into round propagation mineral wool cubes or perlite, and were kept in Gallenkamp growth room. 4ppm of sodium arsenite was given to the As treated group (Figure 2.6)

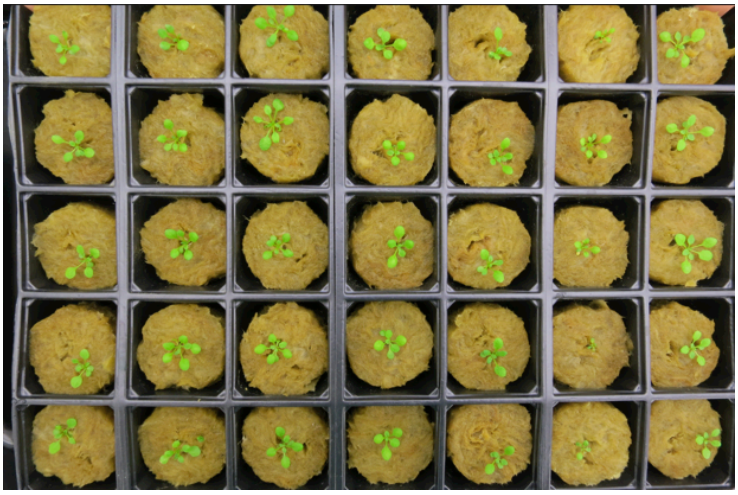

Figure 2.6. A representative image of *Arabidopsis* hydroponic cultivation.

Some minor changes were made in hydroponic cultivation of *A. thaliana* to suit different research proposes.

To measure plant heights change under As stress, plants were grown in hydroponic cultivation for 45 days. The mineral wool cube was applied as the growth media for plant growth.

To measure As content in plant tissues, plants were cultivated hydroponically for 40 days. To separate plant root from growth media easily, perlite was used in this experiment.

To measure the change of plant seeds by arsenite stress, the hydroponic culture system was carried out until the plant was mature and dried. The mineral wool cube was applied as the growth media. The ARACONs system (Figure 2.7) is a seed-harvesting device that can automatically collect seeds produced by individual plants during their entire lifetime and was applied to all plants for seed yield assay.

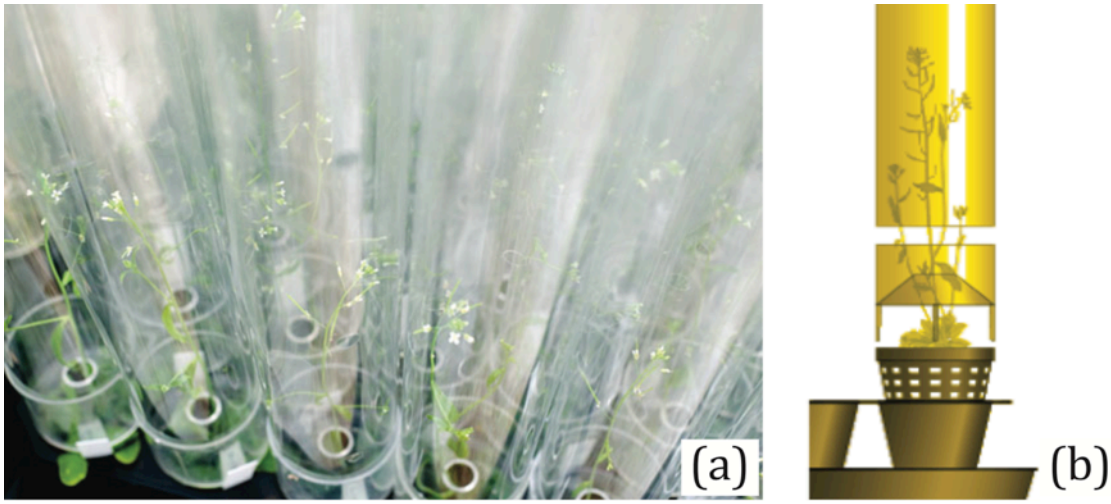

Figure 2.7. ARACONs system a. examples of the ARACONs system; b. Illustrations of the ARACONs system.

2.2.9 Inductively coupled plasma mass spectrometry analysis for *B. juncea* and *A. thaliana*

Brassica seed powder was prepared as described in section 2.2.5 and oil prepared as described in section 2.2.6 as well as *A. thaliana* plant tissue prepared from 2.2.8 were processed with 3% HNO_3 acid digestion procedure (the equipment was washed with 6M HCl) and analyzed using ICP-MS (Ghanthimathi et al. 2012). The inductively coupled plasma mass spectrometry (ICP-MS) analysis of arsenic content was performed by Dr. Matthew Cooper (National Oceanography Centre Southampton).

2.2.10 Analysis of mature *A. thaliana* seed weight and size

After approximately 10 weeks of hydroponic cultivation, mature seeds were harvested from dry siliques that were naturally desiccated, and stored in 1.5mL tubes (Eppendorf) with pierced lids. Total seed weight was measured by weighing all the seeds from 3 individuals of each ecotype cultured hydroponically. Seed weight and size (20 to 30 per ecotype) were measured for all individual plants. Seed weight was measured using a Mettler UMT2 microbalance (Mettler-Toledo, Leicester, UK). Seed size was measured as described in section 2.1.8.

2.2.11 Seed germination assay of *A. thaliana*

All seeds were surface sterilised in 15% (v/v) of the commercial thin bleach (containing 4.5% sodium hypochlorite) for 5min. Then seeds were centrifuged (Eppendorf) at a low speed (2,500 xg) to drain bleach and then washed thoroughly with sterilised $_{dd}H_2O$ for five times during five minutes. 100 seeds of each ecotype were placed evenly into the seeding growth medium in 90mm round Petri dishes (Sterilin) and sealed with micropore tape (3M, UK). Seeds placed in Petri dishes were given a 4^{o}C stratification and then placed in a Gallenkamp growth room for 48h and 72 h. Seeds were recorded for germination under a Leica MZ6 dissecting microscope. 4ppm sodium arsenite was given to the As treated group.

2.2.12 Phenotype measurement assay

All seeds were sterilised with the same approach as described in section 2.1.11. Six seeds of each ecotype were placed evenly into the seeding growth medium in a plant culture container 330ml pot with lid (Greiner Bio-One). Seeds were given a 4^{o}C stratification and then placed in pots and left in Gallenkamp growth room for 2 weeks, in order to measure phenotypic characters changes caused by 4ppm and 6ppm of sodium arsenite.

2.2.13 Rosette measurement of *A. thaliana*

All seeds were sterilised with the same approach as described in section 2.1.11. Six seeds from each ecotype were placed evenly into the seeding growth medium in 140mm round Petri dishes (Sterilin) and the dishes were sealed with micropore tape (3M, UK). Seeds placed in Petri dishes were given a 4^{o}C stratification and then placed in Gallenkamp growth room for 2 weeks. The diameter of a seedling rosette of an individual sample was directly compared and measured.

2.2.14 Post-germination assay of *A. thaliana*

All seeds were sterilised with the same approach as described in section 2.1.11. Two hundred seeds of each ecotype were placed evenly into the seeding growth medium in 90mm round Petri dishes (Sterilin) and the dishes were sealed with micropore tape (3M, UK). Seeds placed in Petri dishes were given a 4^{o}C stratification and the dishes were then placed in Gallenkamp growth room for 7 days. The 6 healthiest seedlings with 2 cotyledons were selected and transferred into 140mm round Petri dish (Sterilin) containing 80mL of seedling growth medium. These Petri dishes were then sealed with micropore tape (3M, UK) and the seeds were cultured for another 7 days. The 140mm round Petri dishes were divided into 24 units, and 6 seedlings of one ecotype were placed in one unit. 4ppm or 6ppm of sodium arsenite was given as an As stress. Seedlings for each ecotype were placed in the same pattern in both the As-treated Petri dishes and the control Petri dishes.

After a 14-day period of growth, a plant shoot was cut from the bottom of hypocotyl and put into a pre-weighed 1.5mL eppendorf tube (Eppendorf). The weight of the shoot and tube was measured using a Mettler UMT2 microbalance (Mettler-Toledo, Leicester, UK).

2.2.15 Chlorophyll content analysis of *A. thaliana* seedings

The chlorophyll in the plant shoot collected from 2.2.14. was extracted with an acetone based chlorophyll extraction assay (Porra 2005; Shibghatallah et al 2013; Ritchie 2006). The plant shoot was transferred into a 15mL falcon tube (FisherScientific) that was wrapped with foil to protect from lights. 5mL 80% acetone was added into each tube and tubes were left at 50^{o}C (Unitemp incubator) overnight until the plant shoot turned white. The absorption of each sample was measured under 663 and 646nm by a WPA Lightwave S2000 UV/Vis Spectrophotometer (BrandTech), after the debris was pelleted by centrifugation (Eppendorf) at a

maximum speed of 20 000 x g in a bench-top centrifuge.

Equations for chlorophyll concentrations (μg/mL)

[Chla] = 12.25 (A_{663}) - 2.55 (A_{646}) (1)

[Chlb] = 20.31 (A_{646}) - 4.91 (A_{663})(2)

[Chlt] = [Chla] +[Chlb] = 17.76 (A_{646}) + 7.34 (A_{663}) (3)

2.2.16 Anthocyanin content analysis of *A. thaliana* seedling

The total anthocyanin content in the plant shoot material collected collected as decribed in 2.2.14 was measured using the pH differential method (Sondheimer and Kertesz 1948; Lee and Wrolstad 2008.). 1% HCl (w/w) in ethanol was used to extract anthocyanin. The plant shoot was transferred into a 15mL falcon tube (FisherScientific) wrapped with foil to prevent lights. 3mL of acid-ethanol was added into each tube and the tubes, which were left at 50^{o}C (Unitemp incubator) overnight until plant shoot turned white. Tubes with sample were centrifuged (Eppendorf) at a maximum speed of 20,000 x g in a bench-top centrifuge in order to separate the supernatant from the remaining tissue. 0.1mL of the extraction was taken carefully from the tube and then transferred into a new tube before mixing with 0.4mL pH1.0 buffer (0.2 mol/l KCl:0.2mol/l HCl:H2O =50:97:53 (v:v:v)). Another 0.1mL of extraction was transferred into a new tube filled with 0.4mL pH4.5 buffer (adding 19.294g NaAc.3H2O and 24mL glacial acetic acid to 500mL total). After the debris was pelleted by centrifugation (Eppendorf) at a maximum speed of 20,000 x g in a bench-top centrifuge, the absorption of each sample was measured using WPA Lightwave S2000 UV/Vis Spectrophotometer (BrandTech) under 520 and 700 nm.

Equations for Anthocyanin concentrations (μg/mL):

A =(A_{520}-A_{700})$pH_{1.0}$-(A_{520}-A_{700})$pH_{4.5}$

[Anthocyanin]= 16.7 x A

The molecular weight of anthocyanin was equivalent to 449.2 g/mol, which is the molecular weight of cyanidin-3-glucoside, since cyanidin-3-glucoside is the most abundant in nature (Francis 1982).

2.2.17 DNA extraction

1. DNA extraction for indel marker selection

Plant DNA extraction for gene expressing examination was effected according to a modified fast micro DNA extraction protocol. The method was applied by placing two small, young leaves in a 1.5mL eppendorf tube (Eppendorf) on ice with 100µl extraction buffer (140mM d-Sorbitol, 220 mMTris-HCl pH 8.0, 800mMNaCl, 22mM EDTA pH 8.0, 0.8% CTAB and 0.1% n-Laurylsarcosine) added. A bench drill with a sterile plastic pestle was used to grind the leaf material to a fine slurry. Sample was then incubated at 65^{o}C for 5min followed by vortexing for 2min with 100µl of chloroform. The debris was pelleted by centrifugation at a maximum speed of 20,000 x g in a bench-top centrifuge for 5min. The supernatant was decanted into a new tube and then precipitated by adding 100µl of iso-propanol. The tubes were incubated at room temperature during 15min. DNA pellets were obtained by centrifugation at maximum speed of 20,000 x g for 20 minutes. The pellet was washed three times with 70% ethanol, and left to dry in a laminar flow hood for 5 minutes or until all traces of the liquid had evaporated. The pellet was dissolved in 80µl sterile Milli-Q H_2O and the sample was stored at 4°C for future use (within 2-3 months).

2. DNA extraction for SNP marker genotyping

Plant DNA extraction for genotyping was done by using QIAampDNeasy Plant Mini Kit (Qiagen, USA) to reach the quality required for Time-of-flight Mass Spectrometry (MALDI-TOF MS) genotyping (Storm and Darnhofer-Patel 2003; Pusch et al. 2002; Jurinke et al. 2004). DNAs from samples were quantified using NanoVue Plus (GE Healthcare Life Sciences) at absorption settings of 260 and 280 nm.

DNA purity (A_{260}/A_{280}) = (A_{260}– A_{320}) /(A_{280}– A_{320})

To confirm DNA quality, DNA purity (A_{260}/A_{280}) of all DNA samples was above 1.5.

2.2.18 DNA amplification for gene expression analysis

For every 20µl reaction, 2µl of extracted DNA solution, 10µl DreamTaq green PCR master mix (2x) (Thermo), 2µl of each forward and reverse primer, and 4µl nuclease-free water were mixed. All reactions were carried out in a G-Storm GS1 Premium Gradient Thermal Cycler (Life science Research, USA). The amplification programmer consisted of: Initial denaturation at 94°C for 4.0 min, followed by 15-35 cycles of denaturation at 94°C for 30 seconds (dependent on experiments), annealing at X°C (where X depends on the primer pairs used) for 30 seconds and extension at 72°C for ~ 1 min (depending on the fragment length, 1 min/kb). A final extension for 10min at 72°C was done to finish the PCR products. Details of PCR primer sequences used for genotyping are shown in (Table 2.3).

2.2.19 Agarose gel electrophoresis

Agarose gels were made by dissolving electrophoresis grade agarose (Invitrogen, UK) or molecular screening agarose (Roche, UK) in 1x TAE buffer (40 mMTris, 1 mM EDTA, glacial acetic acid (1.14%), pH 7.6), and adding ethidium bromide to a final concentration of 0.2 µg/mL before setting. 1% agarose gel was prepared using microwave, 4% agarose gel was prepared by 95^{o}C water bath. The gels were left to set for 30-60 min, and then submerged in 1x TAE buffer containing ethedium bromide in a gel Sub-Cell tank (Bio-Rad). 100bp DNA ladder (New England BioLabs) was loaded on each gel as a molecular weight marker. Electrophoresis was normally carried out using a Powerpac 300 power supply (Bio-Rad) at ~ 140 V. The DNA bands were visualised on a UV transilluminator (UVP, USA).

2.2.20 RT-PCR analysis

RT-PCR analysis was designed for the expression of gene AtPCs1.

RNA extraction:

Two young leaves of *Arabidopsis* Col-0 ecotype were snap frozen in liquid nitrogen and then were ground using a ceramic pestle and mortar. RNA was isolated using PureLink RNA Mini Kit (Ambion, ThermoFisher Scientific) according to the manufacturer's instructions.

First strand cDNA synthesis and PCR

cDNA was synthesised from 1µg total RNA using the ThermoScript RT-PCR system (Invitrogen, UK) according to the manufacturer's instructions.

PCR was then carried out using the cDNA as a template. To analyze the expression level of AtPCs1, a pair of primers of a housekeeping gene was used as internal control for a semi-quality PCR. Each 20µl reaction contained the same regents as described in 2.2.16. The PCR amplification programmer consisted of: Initial denaturation at 94°C for 4.0 min, followed by 20 cycles of denaturation at 94°C for 30 seconds, annealing at 54^{o}C for 30 seconds and extension at 72°C for 1 min. A final extension was performed at for 10 min at 72°C for 10 min to finish extending all PCR products. The PCR products were then run on a 1% agarose gel at 120V for 35 min with 100bp DNA ladder (Invitrogen, UK) as a marker.

2.2.21 Analysis of T-DNA insertion plant line

The *A. thaliana* T-DNA insertion lines in At2g21045 were grown under growth room conditions. Genomic DNA was extracted from the leaves of 2-3 week-old plants. DNA extraction was performed as described in 2.2.15.b. The extracted genomic DNA from each T-DNA insertion lines was used for PCR to identify the genotype of each individual plant.

Two PCR reactions were performed for each plant. Each 20µl reaction mix contained 2µl of extracted DNA solution, 10µl DreamTaq green PCR master mix (2x) (Thermo), 2µl of each forward and reverse primer and 4µl nuclease-free water. The primer

designed for At2g21045 was used in the first reaction; 2μl of T-DNA left border and the reverse primer of At2g21045 were used in the second reaction. The programme of DNA amplification and gel electrophoresis was the same as described in section 2.2.16 and 2.2.17.

2.2.22 QTL mapping of Kas-2 x Ler-0 RI population

QTL analysis was carried out using interval mapping (IM) and composite interval mapping (CIM) implemented in the software Windows QTL Cartographer 2.5. Log-likelihood (LOD) thresholds were used to determine the significance of the QTLs. The analyzes of the means and the principal component scores of the relative and absolute value of arsenite tolerance of 154 lines of Kas-2 x Ler-0 RI population were performed.

2.2.23 Fine mapping analysis

1. Indel marker analysis

Six indel markers were designed as described in section 2.1.5.1. Genomic DNA of both Ler-0 genotype and Kas-2 genotype was used first to test primers. DNA extraction approach was as described in 2.2.17.1, the program for DNA amplification and gel electrophoresis was the same as described in section 2.2.16 and 2.2.17, except that 4% molecular screening agarose (Roche, UK) gel was used.

2. SNP marker analysis

Thirty SNP markers designed from section 2.1.6 and 104 DNA samples prepared from 2.2.17.2 were sent to BGI (Hong Kong, China) for MALDI-TOF MS genotyping. The genotyping data was used to construct a genetic linkage map for fine mapping analysis.

Linkage analysis (to calculate the recombination frequencies between markers,

establish linkage groups, estimate a map distances, and determine a map order) was performed using JoinMap 4.1 software (Kyazam, Netherlands).

Both the linkage map and the physical map of the fine mapping QTL analysis were carried out using interval mapping (IM) and composite interval mapping (CIM) implemented by the software Windows QTL Cartographer 2.5. Log-likelihood (LOD) thresholds were used to determine the significance of the QTLs. The analyzes of the means and principal component scores of the relative and absolute value of arsenite tolerance of the 104 lines of Kas-2 x Ler-0 RI population were performed.

2.2.24 GWAS analysis

GWAS analysis was used for the arsenite tolerance and chlorophyll content data. The analysis was performed with 39 ecotypes and 68,867 gene-based SNPs with trait measures for arsenite tolerance, the amount of total chlorophyll, chlorophyll a and chlorophyll b. The software GenAbel was used because it can filter SNPs according to different quality criteria and correct population structure in various different ways. The GWAS analysis was performed with the help and guidance of Dr. Brennan (Durham University).

2.2.25 Searching database for arsenite tolerance related genes

Genes of the targeted region obtained from section 2.2.23 and 2.2.24 were first collected from Salk Arabidopsis 1,001 Genomes website. The nucleotide and amino acid sequenc information can be compared between accessions in the website. The published transcription analysis data could be obtained using Arabidopsis eFP Browser (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi).

If the function of a gene is known, information of it could be checked from TAIR (http://www.arabidopsis.org/index.jsp).

Chapter 3

Growth of *B. juncea* and *A. thaliana* in the presence of arsenite

3.1 Introduction

In this chapter, the variation of root elongation of *B. juncea* ecotypes under different arsenite stress was first investigated. The As content in hydroponically cultured *B. juncea* seeds and its seed oil were also measured. These two sets of experiments could reveal the arsenite tolerance ability in *B. juncea* and whether this crop can provide a safe oil product when grown in an As polluted environment. A study of phenotypic changes in *A. thaliana* seedlings was also conducted under a range of arsenite concentrations. Finally, in order to study the impact of arsenite on yield changes, the size and weight of *A. thaliana* seeds collected from hydroponic cultivations in the presence of arsenite, as well as the germination rate of these seeds were analyzed. The phenotype change response to arsenite stress of *B. juncea* and *A. thaliana* were fairly similar, which allowed us to restrict in depth study of arsenite tolerance to *A. thaliana* with a view to applying knowledge gained in this model plant to the improvement of *B. juncea* arsenite tolerance (Barnes 2002).

3.1.1 Impact of heavy metal stress on phenotype

At the early stage of this research, the first task was to examine the arsenite tolerance ability in *B. juncea* and the variation in arsenite tolerance among different cultivars. This research was interested in studying the phenotypic changes caused by arsenite stress. Previous researches that include studies on As and other heavy metals can be useful to reference. It has been reported that heavy metals could inhibit the growth and productivity of crops from many studies. Ahmad et al. (2015) found that plants after Cd treatment showed a substantial decrease in height, root length, dry weight, pigments and protein content. Further research on the impact of Cd and Pb in *B. juncea* found that a decline in root length ranged from 17% (with 300 μM Cd) to 54% (with 900 μM Cd) at the pre-flowering stage, as well as a 35% decrease of chlorophyll a (with 900 μM Cd) (John et al. 2012). A research using two-dimensional electrophoresis and digital image analysis examined the changes in differentially

displayed proteins under 24h of As^V in maize seedling root. Among all up- or down-regulated maize root proteins by As^V, a major and functionally homogeneous group of seven enzymes involved in cellular homeostasis for redox perturbation, and four additional functionally heterogeneous proteins was identified by comparing their peptide mass fingerprints against protein- and expressed sequence tag-databases (Requejo and Tena 2005). However, very little is known about the naturally occurring intraspecific variation of As tolerance in plants. When different genotypes of same species showed phenotypic variation, issues like whether this variation is caused by the regulation of the same mechanism, and whether the tolerance mechanisms is As specific mechanism or for general abiotic stress, are all remain unclear. To improve the understating of these questions, research can start from studying the arsenite tolerance variation among various *A. thaliana* genotypes.

3.1.2 Arsenic concentration analysis

Examining the As level in seed oil from plants grown in As contaminated environment was also conducted in this chapter. *B. juncea* seed oil is the major edible oil in India, whereas the residual part of the seeds is used as cattle feed and fertiliser (Ahmad et al. 2015). However, little is known about the As concentration in mustard seed, which is a potential health problem when the crop is grown on As soils. The question related to As content in seed oil need to be answered, and there are several analytical methods that are in principle capable of detecting As in plant tissues. Inductively coupled plasma mass spectrometry (ICP-MS) has been used to analyze As because of its sensitivity in detecting metals as well as several non-metals (Gómez-Ariza et al. 2000; Montes-Bayón et al. 2004). Its sensitivity is as low as one part in 10^{15} parts (part per quadrillion, ppq) on non-interfered low-background isotopes, which is the general requirement for using ICP-MS analysis in high sensitivity (Gómez-Ariza et al. 2000). Atomic absorption spectrophotometry (AAS) is another approach capable of measuring As concentration. One study used AAS to analyze As content in 10ppm and 30ppm As(III) treated *B. juncea* seed; however they were unable to detect As in

the vegetative tissues suggesting the As level was below detection limits (Sinha et al. 2010). Combining the information, this research chose to use ICP-MS to examine the As content in mustard seed oil.

3.1.3 Arsenite stress under hydroponic cultivation

To study the variation of arsenite tolerance in this research, the best growth treatment should avoid using soil. Arsenic (As) speciation in the soil environment is complex (Meharg and Hartley-Whitaker 2002). The toxicity of As depends on the relative distribution of its chemical forms, solubility and bioavailability (Sinha et al. 2010). The soil properties, such as pH, cation-exchange capacity, organic matter, phosphorus concentration and soil redox conditions can affect As speciation. The availability of As in soil also affects plant uptake, and in the end, may determine the suitability or applicability of a particular soil remediation technology (Moreno-Jiménez et al. 2012). The study of the mature plant can therefore use hydroponic system to avoid As complexity in soil. Using media such as rock wool and vermiculite, plus water-soluble nutrients, researchers can regulate the kinds and amounts of nutrients supplied to the plants. Many plants can be grown from seed to seed without ever coming in contact with soil, but it generally requires more intensive care than those grown in soil to insure adequate moisture and to prevent the development of nutrient toxicity or deficiency symptoms (Brenkendaal et al. 1972).

3.2 Materials and Methods

3.2.1 Materials

The detail of 14 ecotypes of *B. juncea* used in this chapter could be found in Section 2.1.1.1;

The information of the 22 genotypes of *A. thaliana* including Bur-0; Ct-1; Col-0; Col-1; Col-3; Edi-0; Hi-0; Kas-2; Kn-0; Ler-0; Mt-0; Nd-1; No-0; Oy-0; Po-0; Rsch-4; Sf-2; St-0; Tsu-0; Wil-2; Wu-0; Zu-0 was given in Section 2.1.1.2.

The hydroponic growth media used for *B. juncea* and *A. thaliana* was given in Section 2.1.3; and the plant seedling growth medium formula was given in Section 2.1.4.

3.2.2 Main methods

The methods used in this chapter included:

Seedling root elongation assay of *B. juncea* and *A. thaliana*;

Hydroponic cultivation of *B. juncea* and *A. thaliana*;

Oil extraction of seeds of *B. juncea*;

ICP-MS analysis;

Mature seed weight and size analysis of *A. thaliana*;

Seeds germination assay of *A. thaliana*.

The detail about these methods were given in Chapter 2.

3.3 Results

The study of arsenite tolerance and accumulation in *B. juncea*

3.3.1. Changes of root length of *B. juncea* under varying arsenite stress

To compare arsenite tolerance among ecotypes of *B. juncea*, this research was started with experiments using the seedling stage. Seeds from all ecotypes were grown under the same growth conditions including the same concentration series of arsenite to observe any morphological differences.

The most common difference observed was in the root length as shown in the example in figure 3.1. Some ecotypes had much longer roots while others were significantly shorter. Those ecotypes, which could elongate much more siginficantlly roots, were categorised as the arsenite tolerance ecotypes, whereas the shorter ecotypes were categorised as the arsenite sensitive ecotypes.

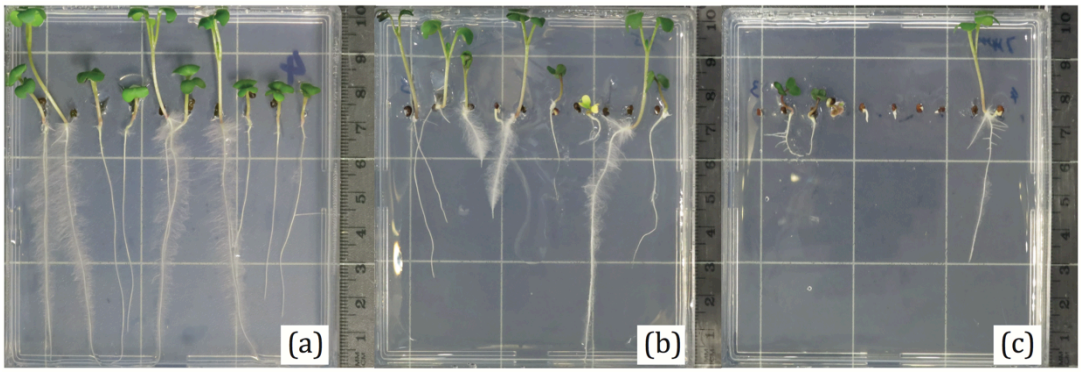

Figure 3.1. Representative images of 8-day old seedlings of *B. juncea* Bra-11 ecotype. a. 0ppm arsenite growth condition. b. 3.5ppm arsenite growth condition. c. 7ppm arsenite growth condition.

Plants have a natural variation in root length. The *B. juncea* cultivars used in this experiment were the Bangladesh local ecotypes and were originally collected from Bangladesh, but their breeding backgrounds were unclear. The root lengths of the arsenite treated plants had the additive effects of both natural variation and arsenite. The mean absolute root length from a total 10 samples of 14 *B. juncea* ecotypes was

calculated as shown in Figure 3.2. Five different sodium arsenite concentrations (1.5ppm, 3,5ppm, 5.0ppm, 7.0ppm and 0ppm as control) were employed in this experiment. The concentration 1.5ppm and 3.5ppm could reflect the level of As in the general contaminated land (Chowdhury et al. 2000). The concentration 5.0ppm and 7.0ppm were selected from a preliminary tests. The concentrations of 10ppm and above were also used for preliminary tests, but 7ppm was the maximum concentration for plants to tolerate. Seeds of *B. juncea* were germinated and grown on the surface of the agar plates and measurements were taken after 8 days of germination (Figure 3.2).

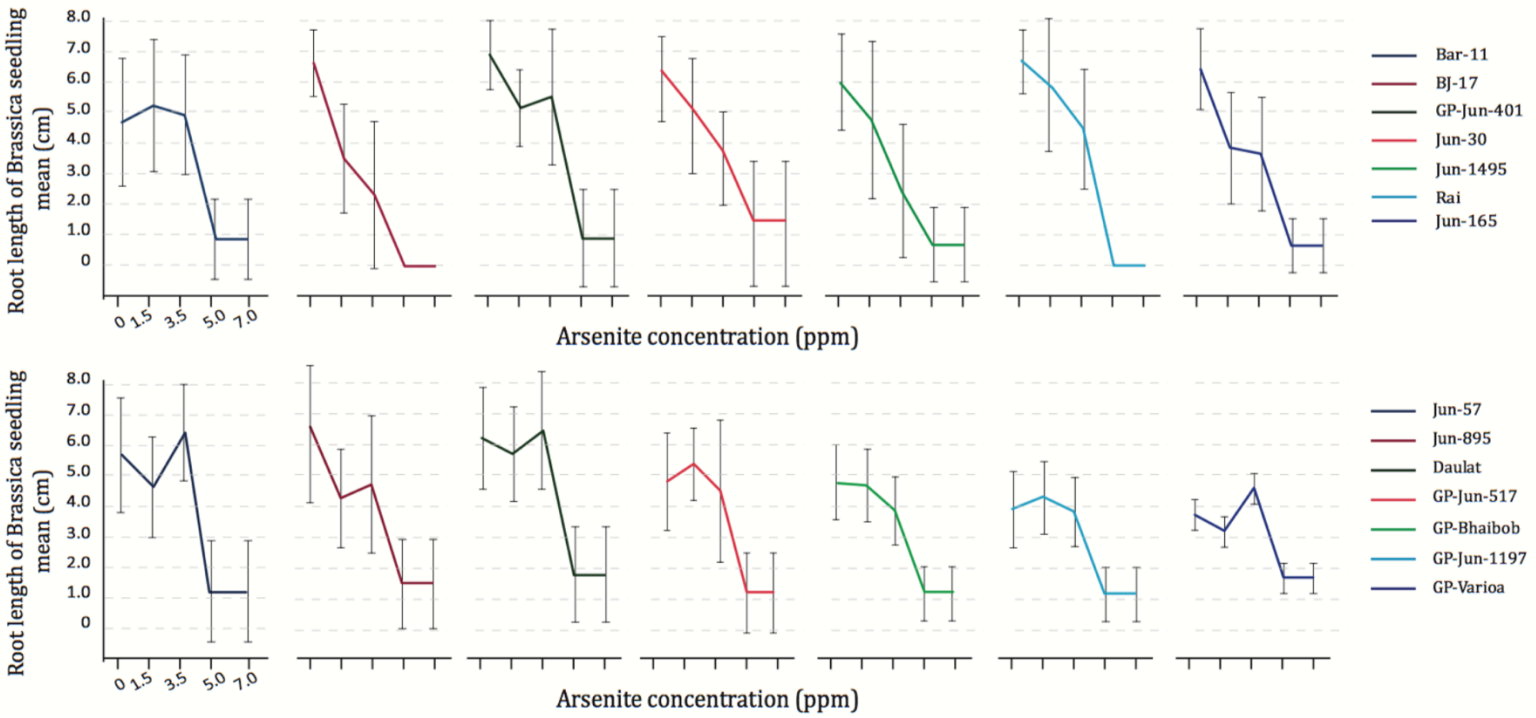

Figure 3.2. The mean root length of 8-day old seedlings of 14 *B. juncea* ecotypes under 1.5ppm, 3.5ppm, 5.0ppm, 7.0ppm arsenite and the control condition (0ppm). Data are means of five replicates (±SD).

Considerable variations amongst seedlings from different ecotypes were observed (Figure 3.2), which suggested the genetic background of ecotypes had a large impact on the root elongation of post-germination stage seedlings (Two-way ANOVA, *F-value*=5.886, *P-value*=0.000). The root length from the same ecotype differed under different arsenite concentrations indicating that different sodium arsenite concentrations had a great impact on the root elongation of the same ecotype (Two-way ANOVA, *F-value* =246.054, *P-value* = 0.000). Both ecotypes and different arsenite concentrations significantly affected root elongation of post-germination stage seedlings (Two-way ANOVA, *F-value* =4.883, *P-value* = 0.000). The adjusted R^2 in this analysis was 64.30%.

In order to reduce the impact of the natural variation of the root length, the relative root length was calculated using the root length of arsenite treated groups divided by control groups. The root length of post-germination seedlings from 14 *B. juncea* ecotypes under different arsenite stress is summarized in Figure 3.3.

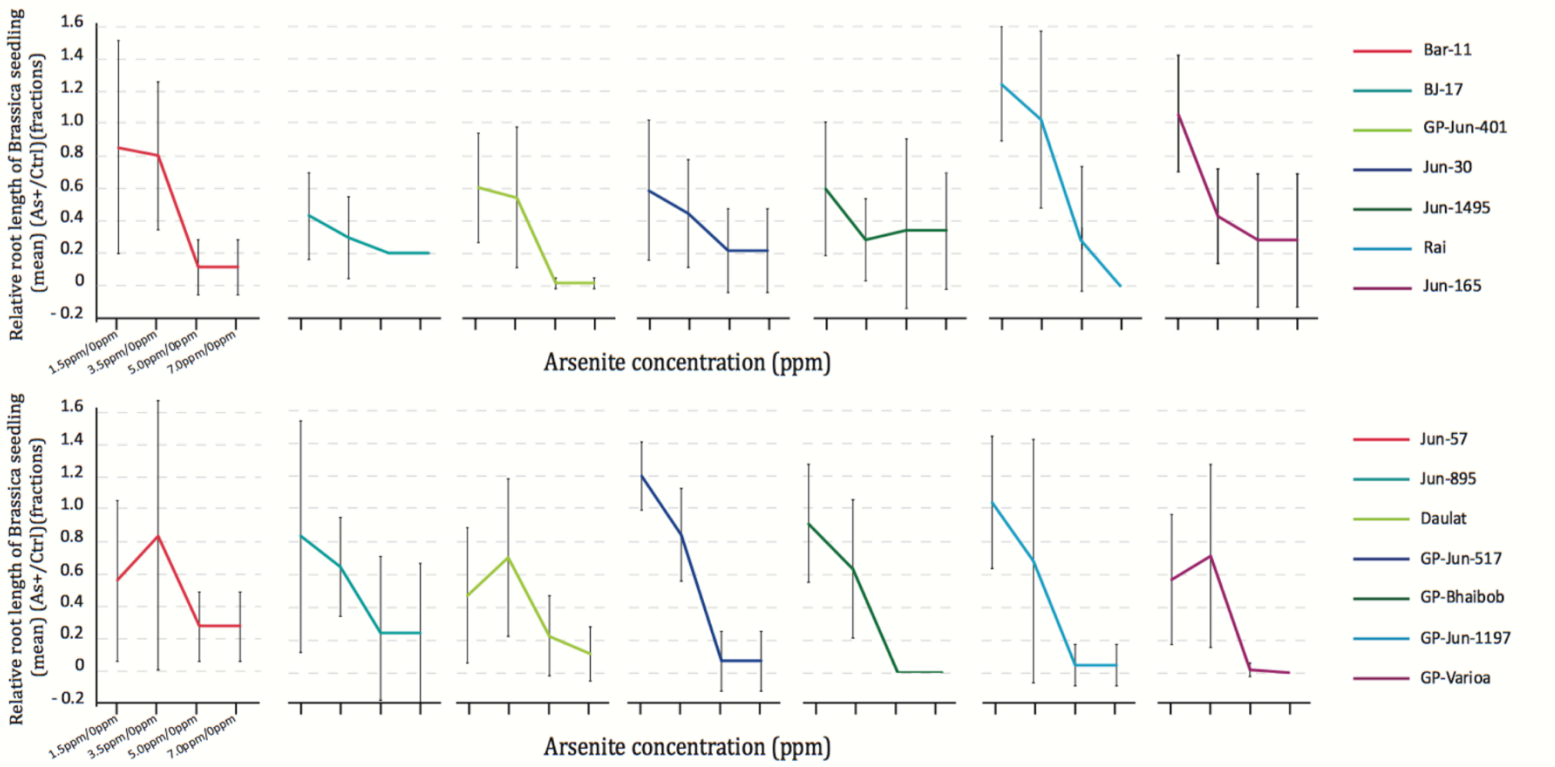

Figure 3.3. The relative root length change caused by arsenite (the root length of arsenite treated seedlings: the root length of the seedlings of the control group). The measurement was from 8-day old seedlings of 14 *B. juncea* ecotypes under 1.5ppm, 3.5ppm, 5.0ppm, 7.0ppm arsenite and the control condition (0ppm). Data are means of five replicates (±SD).

A significant difference was seen concerning the relative root elongation of post-germination stage seedlings from different ecotypes (Figure 3.3; Two-way ANOVA, *F-value* =2.302, *P-value* = 0.006). This result also indicated that different sodium arsenite concentrations had a great impact on relative root elongation of the same genetic background seedlings (Figure 3.3; Two-way ANOVA, *F-value* =41.609, *P-value* = 0.000). However, there was no clear linear correlation between the different ecotypes and arsenite concentrations of post-germination stage seedlings. The adjusted R^2 in this analysis was 22.20%.

The Analysis of variance (ANOVA) statistics analysis showed that arsenite had a significant impact on all seedling root length change. However, a clear observation in Figure 3.3 showed that different ecotypes acted very differently under same arsenite stress, especially the behaviour of root elongation between 0 to 3.5ppm of As. Ecotypes like BJ-17 and Rai, their growth were inhibited when in presence of As. Whereas ecotypes like Bar-11 and GP-Jun-517, their root grew better under 1.5ppm As than the control group. Moreover, ecotypes like Daulat and GP-Jun-401, their root grew better under 3.5ppm As than other lower concentration It was very difficult to summarise a general behaviour of root length change under arsenite stress. To compare the general impact from different arsenite concentration on *B. juncea* seedlings, boxplot graph was used to analyze all *B. juncea* ecotypes under each arsenite concentration.

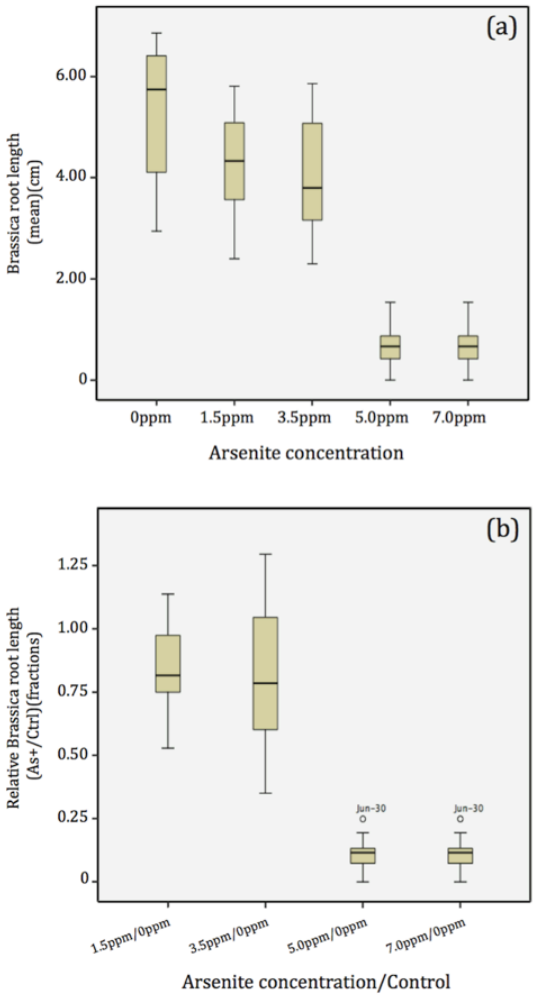

Figure 3.4. a. Boxplot of absolute root length data. b. Boxplot of the change of root length caused by As^{III} (relative root length) under different arsenite concentrations. Data were both collected from all 14 ecotypes of *B. juncea* (± SD).

Between 1.5ppm and 3.5ppm sodium arsenite, a cross-over genotype-environment interaction were observed in different *B. juncea* ecotypes (Figure 3.5).

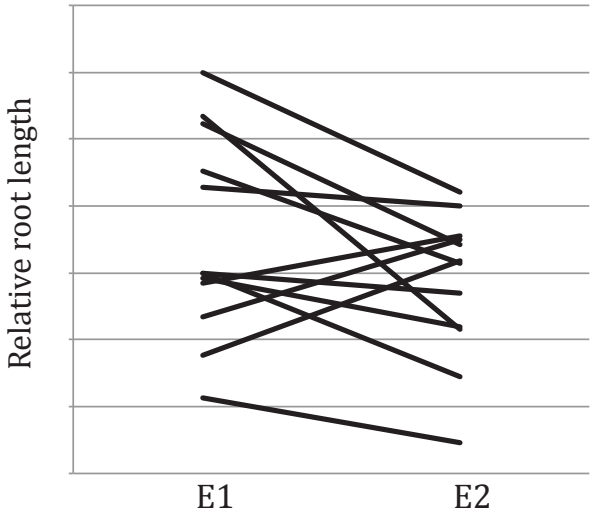

Figure 3.5. The change of root length caused by As^{III} (relative root length) of all tested ecotypes is placed on the axis of two different environments. E1 is the environment 1, which was treated with 1.5ppm sodium arsenite and E2 is the environment 2, which was treated with 3.5ppm sodium arsenite.

It is difficult to summarise a general behaviour of root length change under arsenite stress. Therefore, to compare general arsenite tolerance ability of all *B. juncea* ecotypes, cluster analysis was carried out using mean root length under different arsenite stress to divide all ecotypes into 3 different arsenite tolerance groups (Figure 3.6).

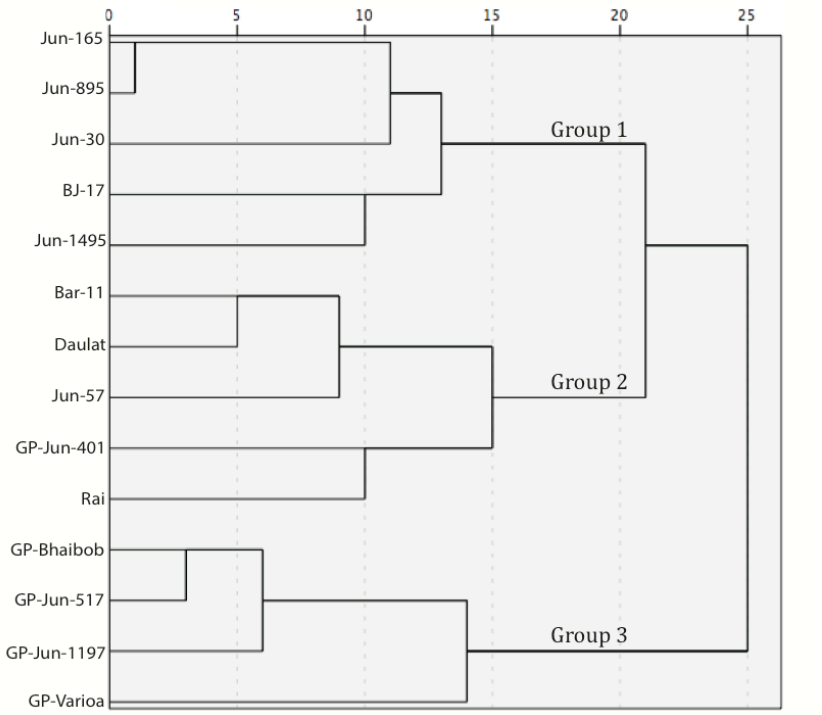

Figure 3.6. Cluster analysis of *B. juncea* root elongation under different arsenite stress. Group 1 has the highest As tolerance; group 2 has the medium As tolerance; group 3 has the lowest As tolerance.

Jun-165, Jun-895, Jun-30, BJ-17 and Jun-1495 were grouped together in Group 1, which was the highest arsenite tolerance ability group. The arsenite tolerance ability of Bar-11, Daulat, Jun-57, GP-Jun-401, and Rai belonged to the middle level Group 2. GP-Bhaibob, GP-Jun-517, GP-Jun-1197 and GP-Varioa were grouped together as they were relatively more sensitive to arsenite stress and put together in Group 3.

3.3.2 Impact of arsenite on hydroponically cultured *B. juncea*

In order to understand the reaction of the mature plant under arsenite stress, two or three *B. juncea* ecotypes from each group (Figure 3.6) were selected for hydroponic culture. In Figure 3.7, measurements were made of the mean heights of 6 hydroponically cultured *B. juncea* ecotypes from both control conditions and 1ppm sodium arsenite treated conditions. By comparing the height of the 60-day old plants, it showed that 1ppm arsenite did not have an significant impact on the height of *B.*

juncea plants (Independent sample T-test, plant height *P-value* =0.05278).

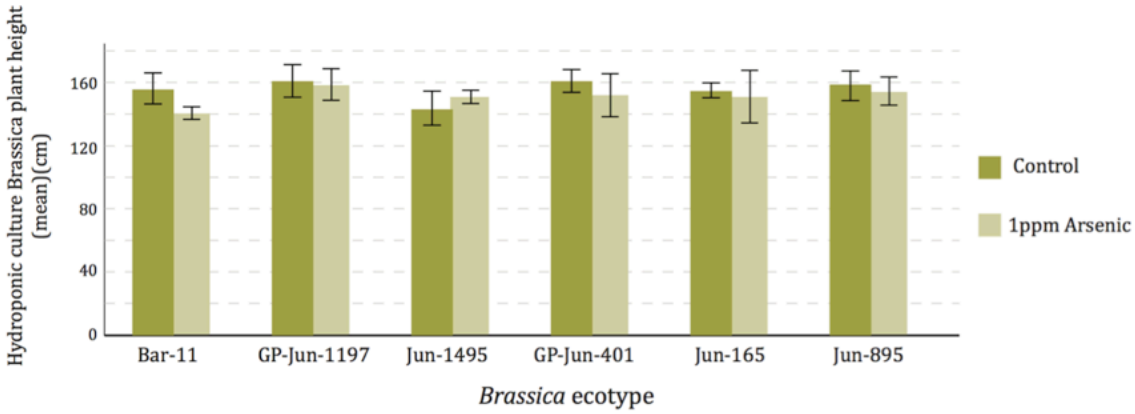

Figure 3.7. The mean height of *B. juncea* plants cultured hydroponically under control conditions and arsenite treatment. Data are means of five replicates (±SD).

Brassica seed oil of 6 ecotypes was extracted using hexane (Figure 3.8). Total seed weight and the oil weight of each hydroponically cultured plant are given in Figure 3.9. The oil yield was calculated from using total oil weight against total seed weight, and oil yield of each *B. juncea* ecotype was also given in Figure 3.9.

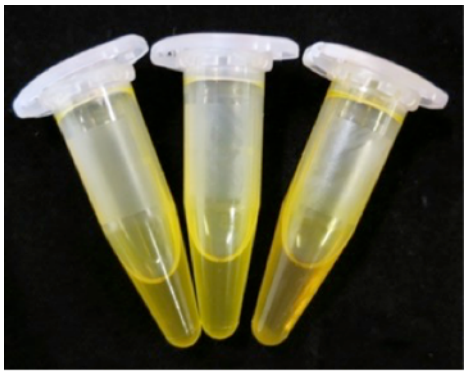

Figure 3.8. An illustration of total seed oil extracted from an individual plant. From left to right are the ecotype Bar-11, Jun-1495 and GP-Jun-401.

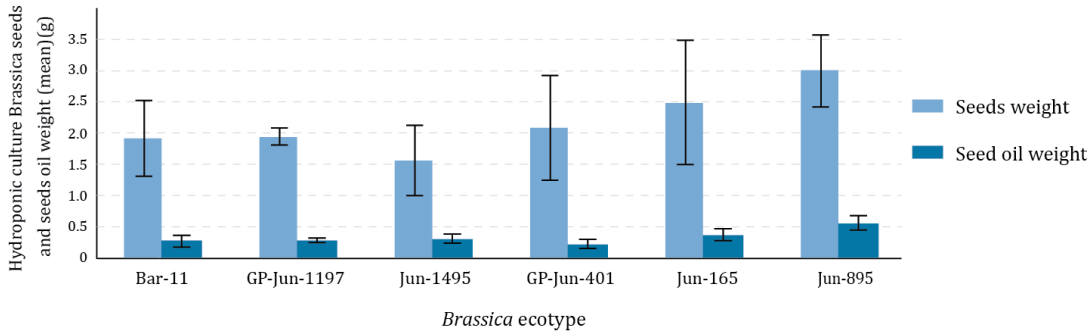

Figure 3.9. The seeds, seed oil and the oil yield from 6 *B. juncea* ecotypes cultured hydroponically under 1ppm arsenite treatment. Data are means of three replicates (± SD).

In the 1ppm arsenite treated group, the total seeds weight of different ecotypes was not significantly different (One-way ANOVA, *F-value* =1.688, *P-value* =0.212), indicating that these *B. juncea* ecotypes have similar seed yield. However, a clear difference of the total seed oil weight extracted from the hydroponically cultivated seeds was observed among ecotypes (Figure 3.9. One-way ANOVA, *F-value* =5.934, *P-value* = 0.005). There was also a notable difference of oil yield among ecotypes after As treatment (One-way ANOVA, *F-value* =3.711, *P-value* = 0.029).

ICP-MS was employed to analyze As concentration of both *B. juncea* seed and seed oil cultured hydroponically. The concentration of As in seeds (Figure 3.10.a) was significantly different among *B. juncea* ecotypes (One-way ANOVA, *F-value*=7.421, *P-value*=0.002). The concentrations of As in seed oil from different *B. juncea* ecotypes (Figure 3.10.b) were also different (One-way ANOVA, *F-value*=26.780, *P-value*=0.000). Using the mean As concentration (C_{As}) in seeds and seed oil of each ecotype (figure 3.10.a and b) and the total seed weight and seed oil weight (Figure 3.9), the total amount of As in seeds and seed oil in every ecotype was calculated (Figure 3.10.c and d)

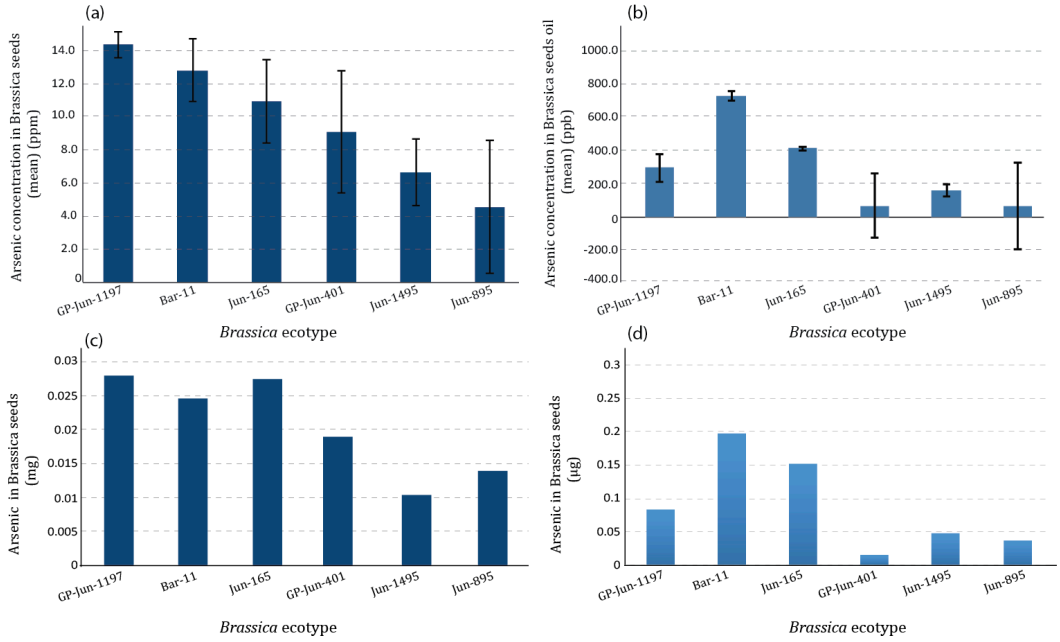

Figure 3.10. a. The mean As concentration (C_{As}) in seeds of 6 *B. juncea* ecotypes in arsenite treatment groups; b. The mean C_{As} in seed oil of 6 *B. juncea* ecotypes in arsenite treatment groups; c. The total amount of As in seeds of 6 *B. juncea* ecotypes in arsenite treatment groups; d. The total amount As in seeds seed oil of 6 *B. juncea* ecotypes in arsenite treatment groups; Data are means of three replicates (±SD).

Comparing Figure 3.10. a and b, it indicates that concentration of As in seeds does not correlate to that in seed oil. GP-Jun-1197 has the highest As concentration in seed powder but not the highest As concentration in seed oil. However, it was clear that, in general, if the seed contains the high concentration of As, the seed oil will contain the relatively high concentration of As too. To check whether the high As concentration in seed oil was caused by a high oil yield or not, the oil yield calculated from Figure 3.9 is compared with “total As in seed oil/ total As in seed” in Table 3.1. It is clear that the high As content in seed oil is not strongly correlated to high oil yield.

Table 3.1 Comparing oil yields of 6 *B. juncea* ecotypes and the As transportation from its seeds to seed oil.

Ecotype	Oil yield (seed oil weight:seed weight)%	As_{oil}: As_{seed} (total As in seed oil: total As in seed)%
Bar-11	14.04%	0.801%
GP-Jun-1197	14.33%	0.297%
GP-Jun-1495	20.18%	0.457%
GP-Jun-401	10.79%	0.078%
J-165	15.48%	0.556%
J-895	15.54%	0.266%

3.4 Results

The study of arsenite tolerance and accumulation of *A. thaliana*

3.4.1 Changes of the root length of *A. thaliana* genotypes under varying arsenite stress

To study the phenotypic characters of arsenite tolerance in *A. thaliana*, seeds from all genotypes were grown at the same time and under the same growth conditions to observe any differences in their features. Since root length was used to compare the As tolerance in *B. juncea*, the tolerance ability of *A. thaliana* was also examined using the root length as shown example in figure 3.11. Ecotypes that could elongate much greater under arsenite stress were categorized as the arsenite tolerance genotypes, whereas those ecotypes could not elongate much were categorized as the arsenite sensitive genotypes.

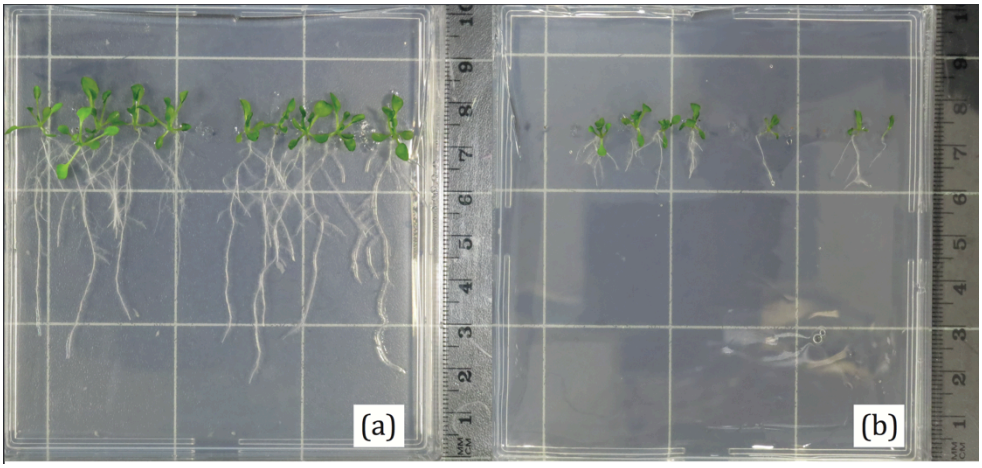

Figure 3.11. Representative images of 14-day old seedlings of *A. thaliana* Col-0 ecotype. a. 0ppm arsenite free group. b. 4ppm arsenite treated group.

Arabidopsis plants showed a variation in the root length in the control group as well as under arsenite stress. To study *Arabidopsis* root elongation under different arsenite concentrations, the mean absolute root length of 10 seedlings selected from each 20

A. thaliana ecotypes was calculated and put into figure 3.12.

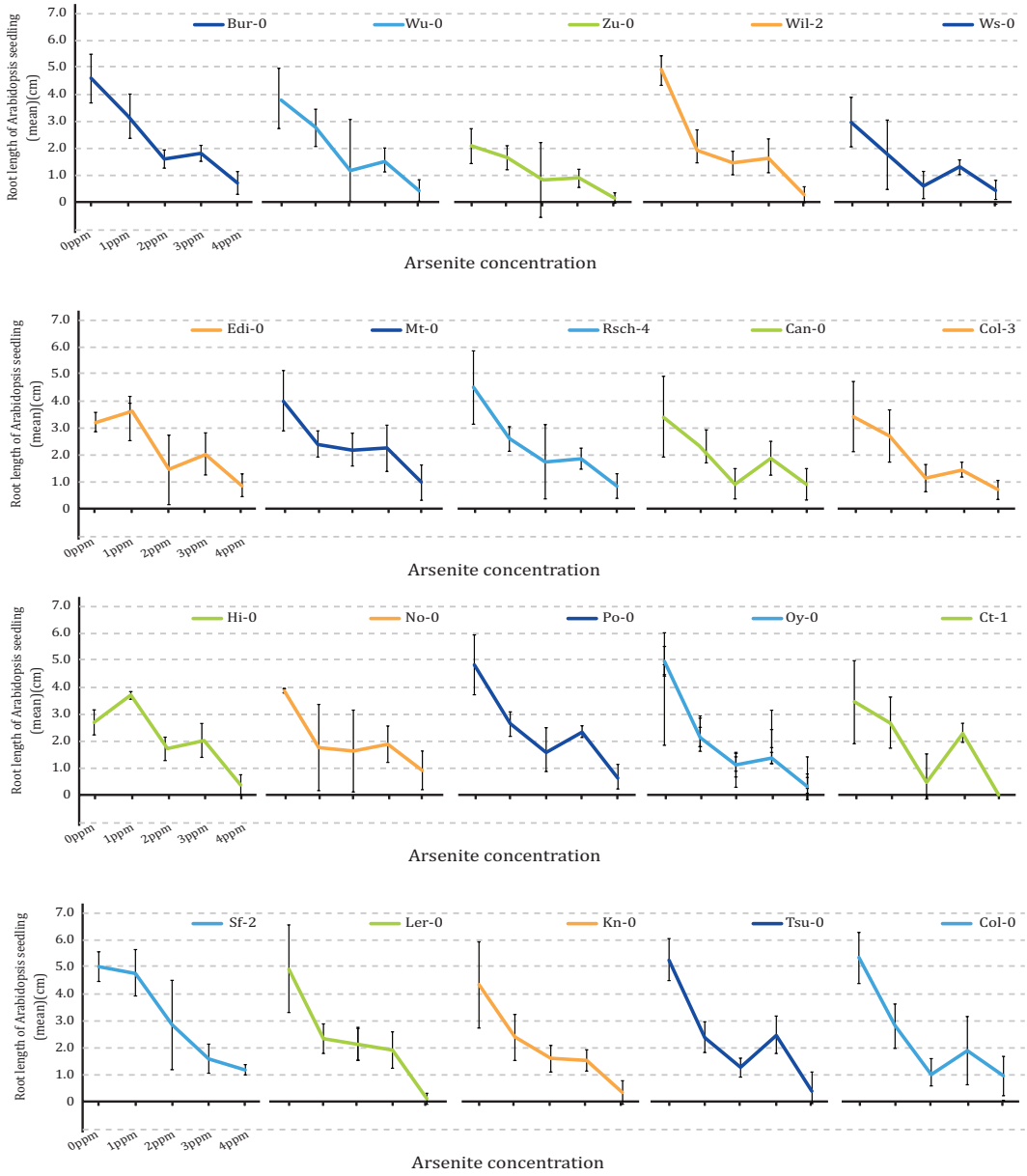

Figure 3.12 The mean root length of 20 *A. thaliana* ecotypes grown under 1ppm, 2ppm, 3ppm and 4ppm arsenite stress. Data are means of four replicates (±SD).

Considerable variations of the root length of 14-days old seedlings from different ecotypes were observed (Figure 3.12; Two-way ANOVA, *F-value*=7.992, *P-value*=0.000). The sodium arsenite concentration had a significant effect on root elongation from the same ecotype seedlings (Two-way ANOVA, *F-value*=408.696, *P-value*=0.000). Both ecotypes and arsenite concentration had a significant impact on

root elongation of seedlings (Two-way ANOVA, *F-value*=3.426, *P-value*=0.000). The adjusted R^2 in this analysis was 67.50%. The adjusted R^2 was used to determine how well the model fits the data. R^2 =Explained variation / Total variation. In general, the higher the R^2, the better the model fits your data. The adjusted R^2 in this analysis was 67.50%, which was not very high. The change of root length caused by As^{III} (relative root lengths) of 20 *A. thaliana* ecotypes under different arsenite concentrations were presented in Figure 3.13.

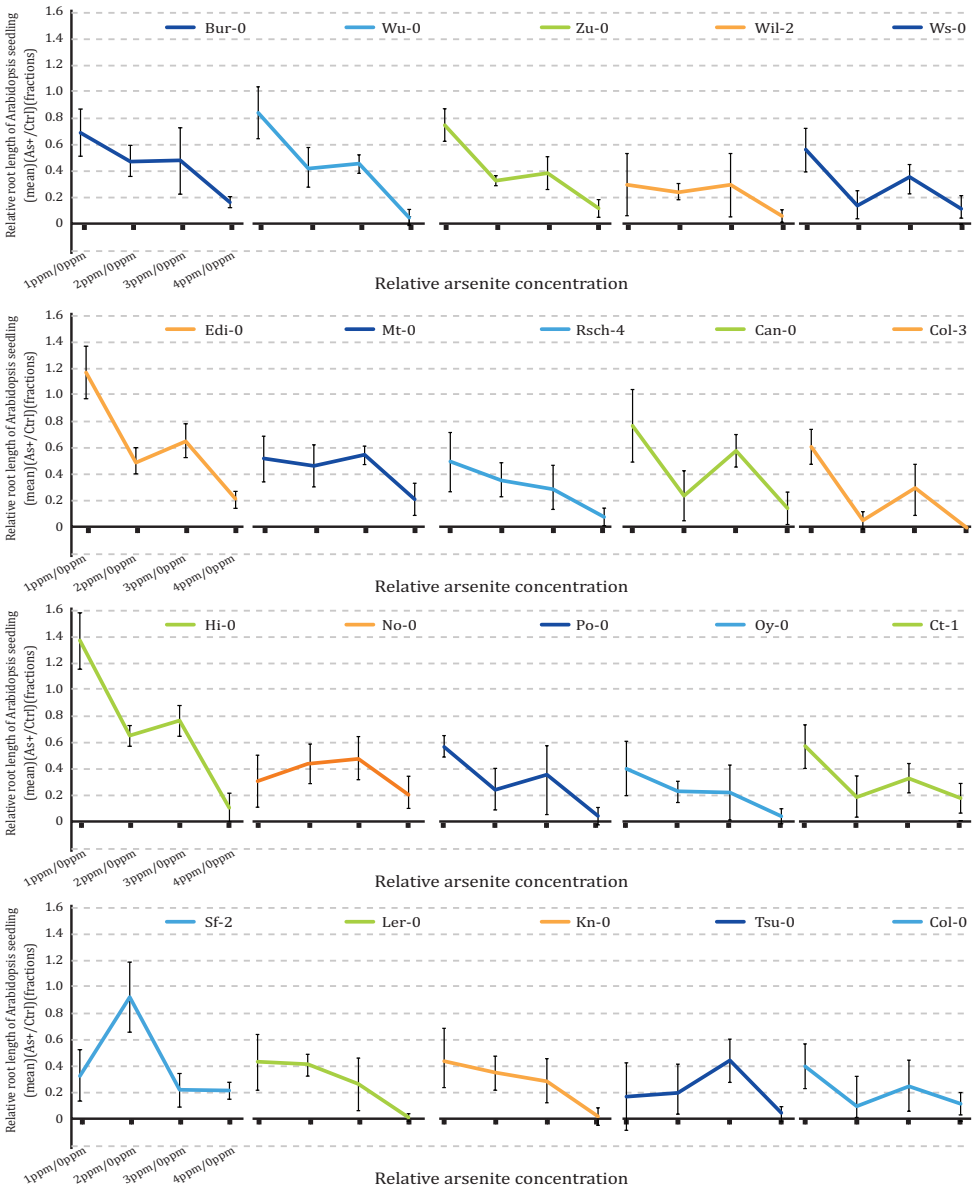

Figure 3.13. The mean change of root length caused by As^{III} (relative root length) of 20 *A. thaliana* genotypes grown under 1ppm, 2ppm, 3ppm and 4ppm arsenite stress. Data are means of four replicates (±SD).

14-day old seedlings from different ecotypes exhibited significant variations in relative root elongation (Figure 3.13; Two-way ANOVA, *F-value*=12.752, *P-value*=0.000). Different sodium arsenite concentrations also have a significant effect on root elongation from the same ecotype seedling (Two-way ANOVA, *F-value*=152.700, *P-value*=0.000). Both ecotypes and arsenite concentration had a notable effect on root elongation of post-germination stage seedlings (Two-way ANOVA, *F-value*=4.574, *P-value*=0.000). The adjusted R^2 in this analysis was 54%.

In the study of *A. thaliana*, it was also very difficult to summarize a general behavior of root length change under arsenite stress and to compare the general impact from the increased arsenite stress to all tested *A. thaliana*. For instance, some genotypes like Ler-0, Sf-2 and Kn-0 heir growth were inhibited when in presence of As. However many genotypes like Wu-0, Ws-0. Can-0, Ct-1, Tsu-0 and Col-0, their root grew better under 3.0ppm As than the lower concentration of As. Hence, a boxplot graph was used to analyze all *A. thaliana* ecotypes under each arsenite concentration. The boxplot of absolute root length data clearly showed a decreasing root length with an increasing arsenite concentration. However, the average root length of all 20 ecotypes from the 3ppm arsenite treated group was longer than the 2ppm arsenite treated group (Figure 3.14). Compared with the relative root length data, the difference between different arsenite concentrations was smaller in the absolute root length group. The relative root length also decreased when arsenite concentration increased.

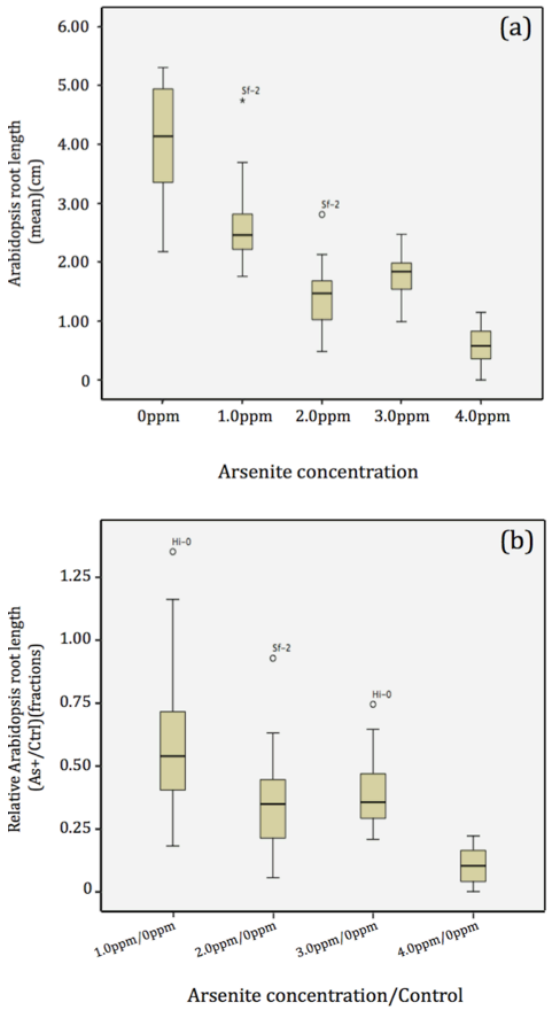

Figure 3.14. a. Boxplot of a. absolute root length. b. Boxplot of the change of root length caused by As^{III} (relative root lengths) under different arsenite concentrations. Data were collected from all 20 ecotypes of *A.thaliana* (± SD).

A general root behaviour pattern can be observed from the boxplot analysis, as the root length elongation was inhibited in presence of As. An exception was also found from 3.0ppm As treatment. The mean root length data calculated from all 3.0ppm As treated genotypes suggested that *A. thaliana* seedlings root grew longer under 3.0ppm of As than the lower concentration. In the analysis of *A. thaliana*, a cross-over genotype-environment interaction was found between 2ppm and 3ppm sodium arsenite (Figure 3.15).

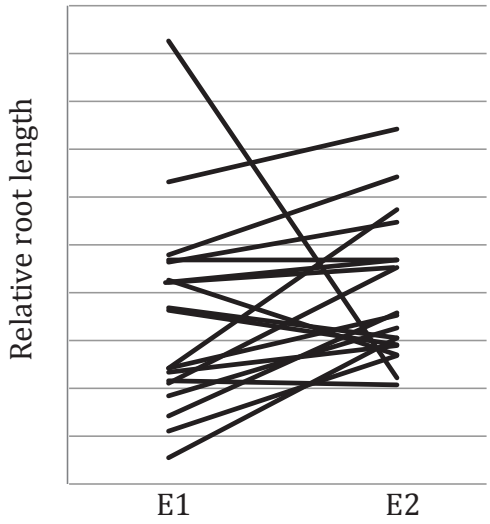

Figure 3.15. The change of root length caused by As^{III} (relative root lengths) of all tested *Arabidopsis* genotypes under two different conditions. E1 represents the environment contains 2.0pm sodium arsenite and E2 represents the environment contains 3.0ppm sodium arsenite.

3.4.2 Impact of arsenite on hydroponically cultured *A. thaliana*

In order to study mature plants behaviour under arsenite stress, a hydroponic cultivation system was employed. Seeds were germinated on a seedling growth medium and transferred on a rockwool cube at day 7. Twenty *A. thaliana* genotypes were grown in Hoagland solution with 4ppm of sodium arsenite. The true leaves number of plants was measured at day 20 and the height of plants was measured at day 45 (Figure 3.16).

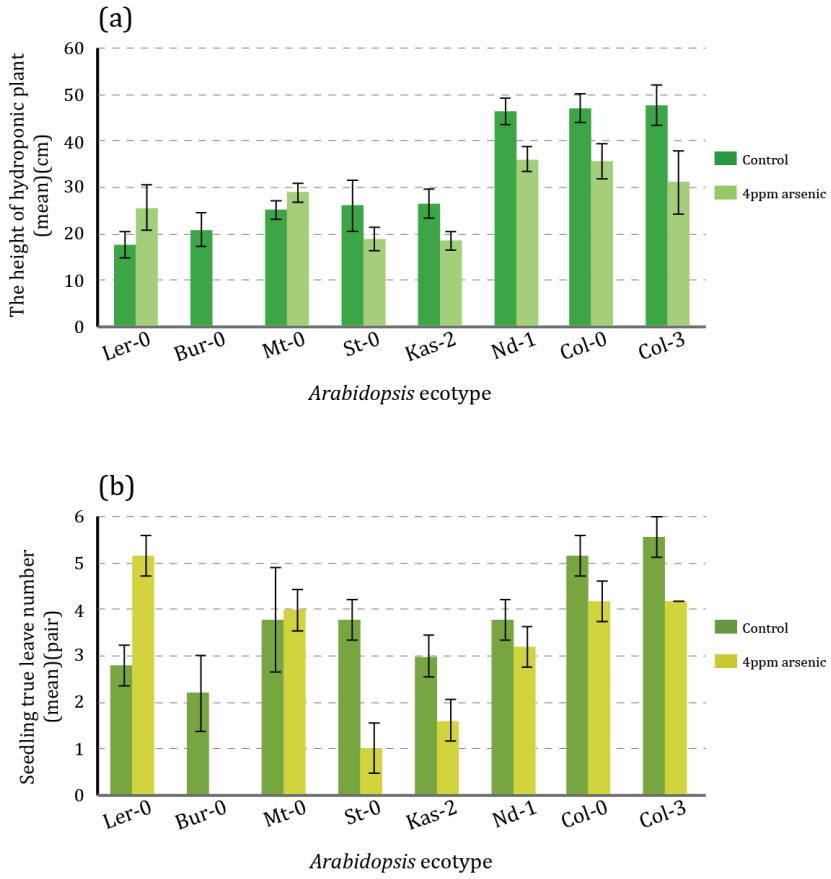

Figure 3.16. a. The mean height of 8 hydroponically cultured *A. thaliana* (45-day old) in control conditions and 4ppm sodium arsenite. Data are means of three replicates (±SD); b. The mean of the number of true leaves of 20-day old seedlings. Data are means of four replicates (±SD).

The height of 45-day old plants was significantly affected by their genetic background (Two-way ANOVA, *F-value*=126.326, *P-value*=0.000), as also by the concentrations of the sodium arsenite (Two-way ANOVA, *F-value*=120.525, *P-value*=0.000). The interaction between both ecotypes and arsenite stress also had a major effect on plant height of 45-day old plants (Two-way ANOVA, *F-value*=22.950, *P-value*=0.000). The adjusted R^2 in this analysis was 93.60%. Genotype Bur-0 could not grow in arsenite treated hydroponic culture; therefore, no data is presented.

The development of plant leaves was also clearly affected by its genetic background

(Two-way ANOVA, *F-value*=67.122, *P-value*=0.000), and by the concentrations of the sodium arsenite (Two-way ANOVA, *F-value*=56.390, *P-value*=0.000). The interaction between both ecotypes and arsenite stress had a significative effect on plant leaf development (Two-way ANOVA, *F-value*=23.095, *P-value*=0.000). The adjusted R^2 in this analysis was 89.70%.

Comparing the true leaf number of 20-day old plants and the height of 45-day plants, a positive correlation was found in the Pearson correlation analysis (the value of R was 0.713 at the significant level of 0.05).

3.4.3 Analysis of the change of seed size and weight under arsenite stress

The total seeds of all individuals were collected using the ARACON system (refer to Section 2.2.8.c). The total seed weights of 22 *A. thaliana* ecotypes cultured hydroponically under the control condition and 4ppm arsenite treatment were measured (Figure 3.17).

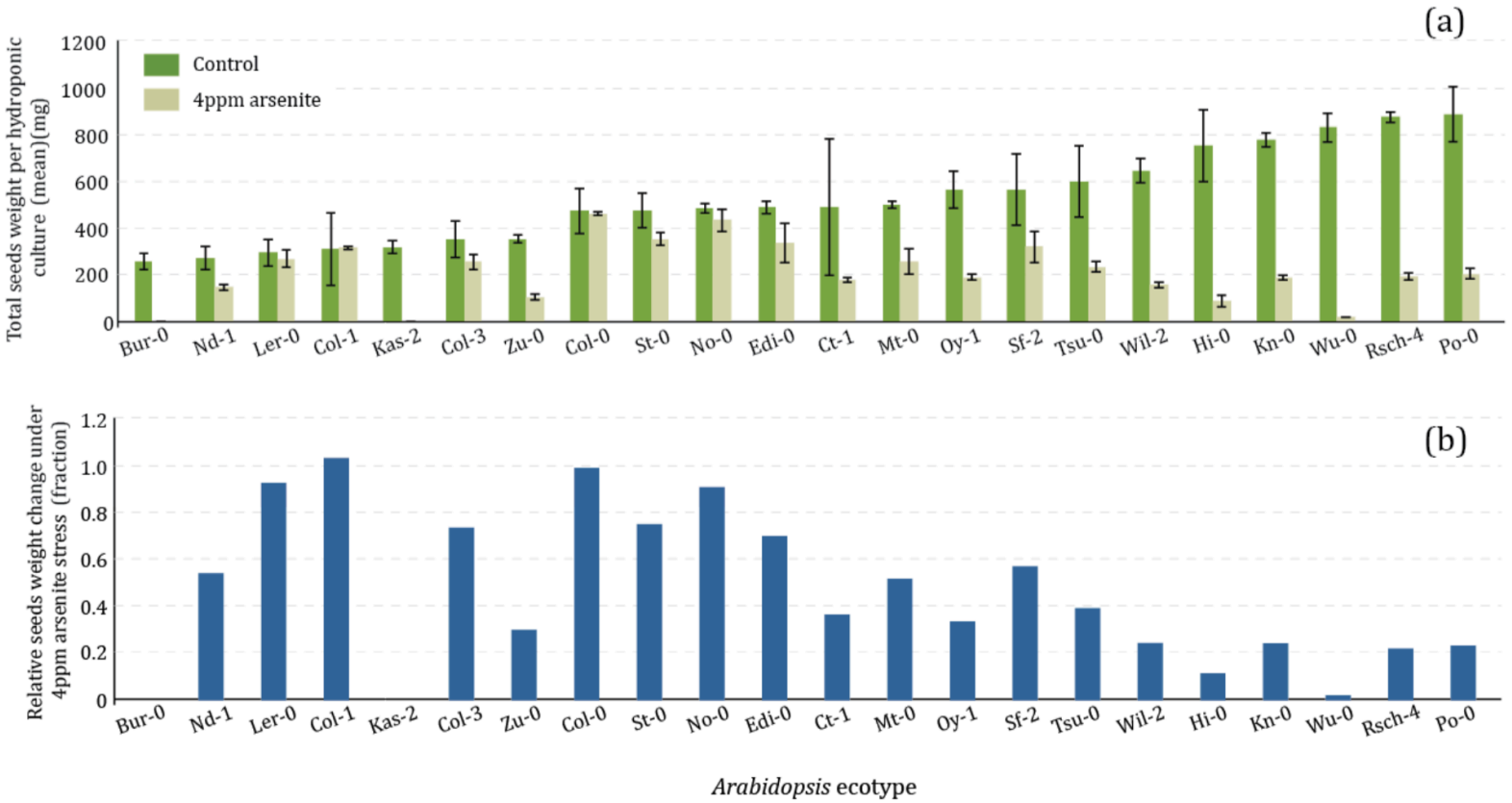

Figure 3.17. a. The mean weight of the total seeds from 22 hydroponically cultured *A. thaliana* ecotypes from the control group and arsenite treatment group. Data are means of four replicates (±SD). b. The relative total seed weight change (weight of seeds of arsenite treated group: weight of seeds of the control group) of 22 hydroponically cultured *A. thaliana* ecotypes.

Total seeds weight was significantly affected by the genetic background (Two-way ANOVA, *F-value*=10.920, *P-value*=0.000), and the sodium arsenite concentrations (Two-way ANOVA, *F-value*=475.303, *P-value*=0.000). There was no significant impact from the interaction between ecotypes and arsenite stress. The adjusted R^2 was 88.10%, so the data analysis explained most of the variability of the response data around its mean. In this experiment, genotype Bur-0 and Kas-2 could not grow and no seeds were produced under arsenite treated hydroponic culture. It appeared that the natural high yield genotypes, such as Po-0, Rsch-4, Wu-0 and Hi-0, their seed yields were highly inhibited under As stress. Other genotypes, which had low yield in control group, their seed yield did not suffer from As stress too much. The seed yield difference between As treated group and the control was less then 10%. The genotypes acted this way were Ler-0, Col-0, No-0 and Col-1.

About twenty seeds from each plant were taken to measure their weight in order to get the pre-seed weight. Images of seeds were taken and analyzed to obtain the seed size data. The change of seed size and weight caused by As^{III} were significantly correlated (R=97.2%, *P-value*=0.000) at the 0.01 level. These data was summarized and given in supplementary data 2.

100 hydroponically cultivated seeds were used for a germination assay, including both control and arsenite treated groups. The number of germinated seeds was counted after 24 hour. In figure 3.18, no data was available for genotypes Bur-0 and Kas-2 due to no seeds produced under hydroponic cultivation with arsenite stress.

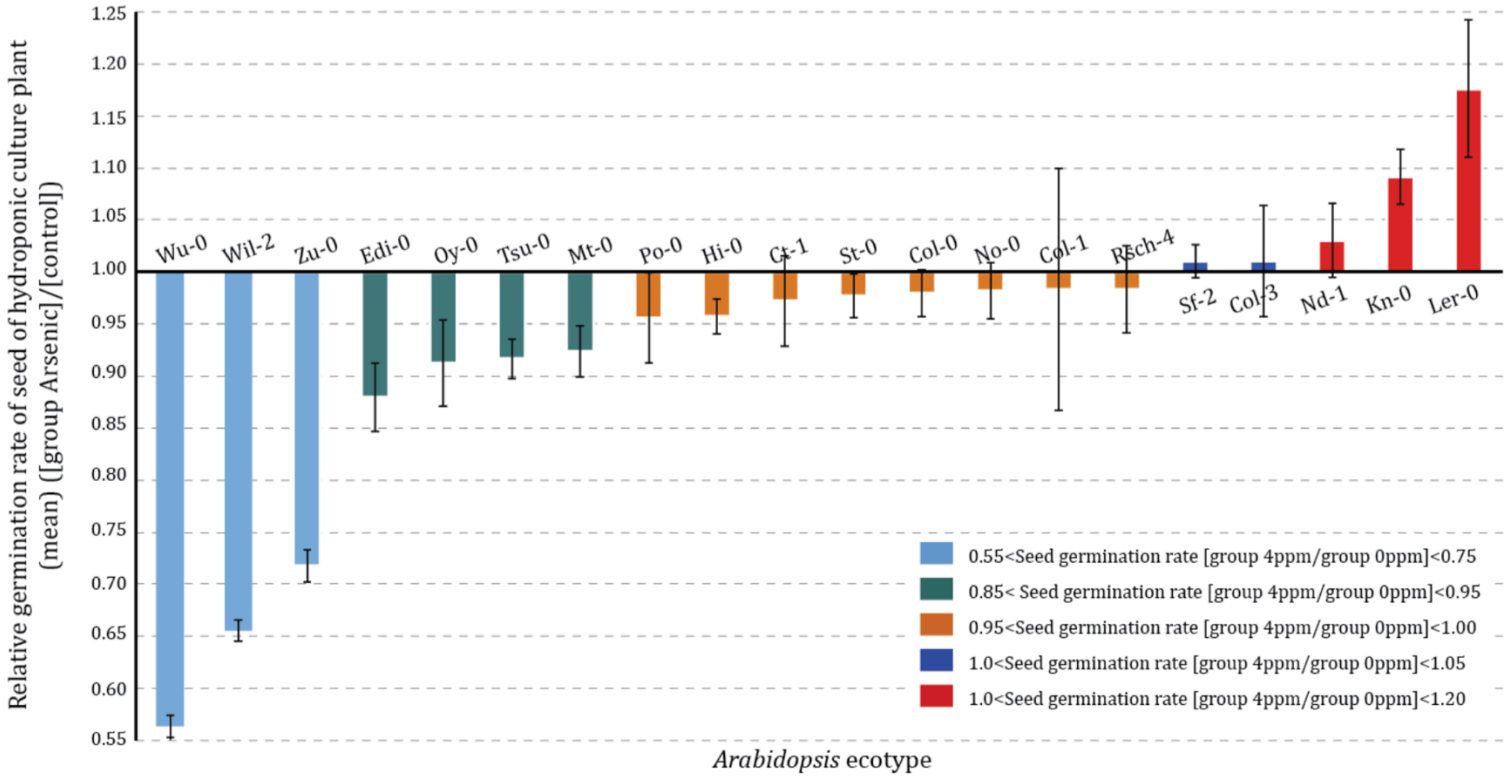

Figure 3.18. The mean relative germination rate of 20 *A. thaliana* ecotypes. The seeds obtained from hydroponically cultivated plants in control and arsenite treatment groups. 100 seeds were used per each germination assay. Data are means of four replicates (±SD).

In Figure 3.18, X axis and Y axis cross at value 1.0. Those arsenite treated groups, which had higher germination rate than the control, are represented above the X-axis, whereas the smaller ones are placed below the X-axis. There was a significant difference between hydroponically cultured plants with 4ppm sodium arsenite treatment and control plants in the germination rate of 100 *Arabidopsis* seeds (Two-way ANOVA, *F-value*=92.682, *P-value*=0.000). The ecotypes also had a significative effect on the germination rate (Two-way ANOVA, *F-value*=22.346, *P-value*=0.000). The interaction between ecotypes and arsenite stress had a signification effect on germination rate (Two-way ANOVA, *F-value*=26.823, *P-value*=0.000). The adjusted R^2 in this analysis was 87.30%. The change of germination rate caused by As^{III} of each ecotype was correlated to both relative seed size (R=52.7%, *P-value* = 0.017) and relative seed weight (R=53.5%, *P-value* = 0.015) at the 0.05 level. Data of *A. thaliana* seed germination obtained from hydroponic cultivated seed were cataloged into five groups. Using the As treated hydroponic cultivar to against control hydroponic cultivar, genotypes Wu-0, Wil-2 and Zu-0 had the lowest ratio, which were lower than 70%, suggested the seed vigor had been depressed the most from As among all tested group. When treated with As, the next generation of these genotypes would have low germination rate. While genotypes like Ler-0, the seed vigor had been increased nearly 20%, which was an interesting discover.

From the seed yield analysis and the seed vigor analysis under As stree, Ler-0 could be confirmed to have the best resistance ability. The total seed yield of genotype Ler-0 was hardly affected by arsenite (Figure 3.17), the size and weight of each seed increased 15% and 25%, respectively, under arsenite stress (supplementary data 2). Seeds collected from arsenite treated groups even had higher germination rate than seed collected from the control group (Figure 3.18).

3.4.4 Analysis of arsenic content in plant tissues

ICP-MS was used to quantify the arsenite concentration in both plant shoots and roots of eight *A. thaliana* genotypes. The mean concentrations of As in both shoots and roots was calculated and given in Figure 3.19 a and b.

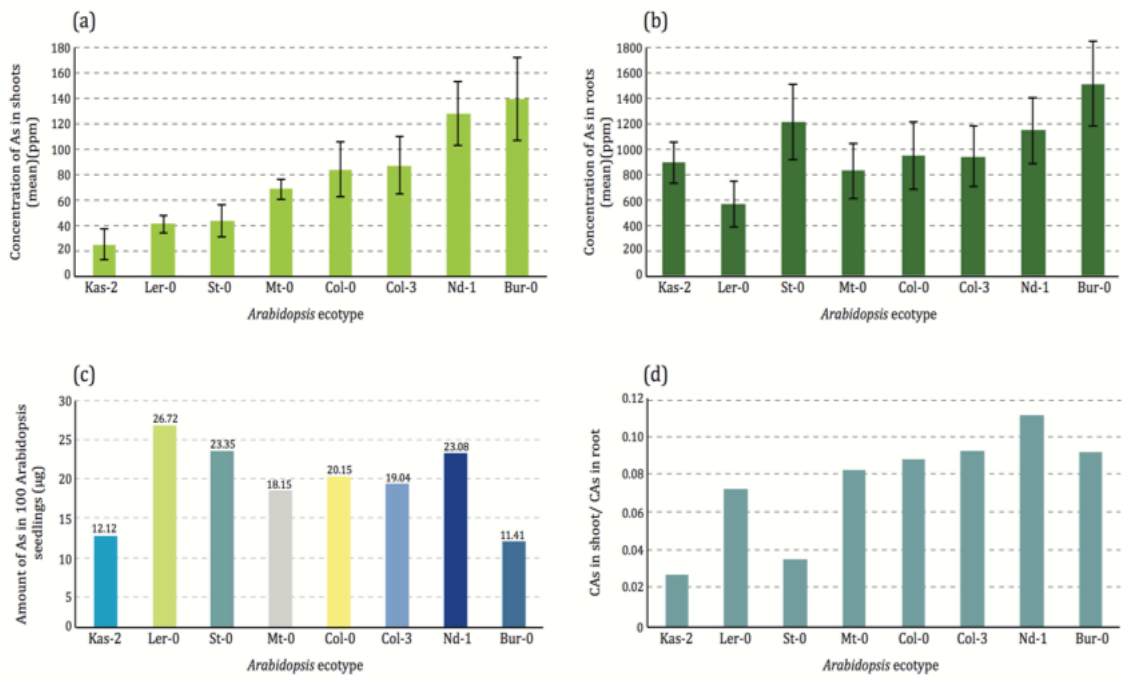

Figure 3.19. a. The mean arsenite concentration (C_{As}) in 40-day old plant shoots. Data are means of three replications (±SD). b. The mean arsenite concentration (C_{As}) in roots. Data are means of three replications (±SD). c. The total As in 100 15-day old *Arabidopsis* seedlings. d. Using the ratio between the arsenite concentration in shoots and roots to represent the As transfer ability of 8 hydroponically cultured *A. thaliana* ecotypes.

The concentrations of arsenite in plant shoots and roots were significantly correlated (R=0.50, *P-value* = 0.000) at 0.05 level. The genetic background of genotypes had a significant effect on arsenite concentration in both plant shoots (One-way ANOVA, *F-value* =17.062, *P-value* = 0.000) and roots (One-way ANOVA, *F-value* =4.176, *P-value* = 0.009). As concentrations in shoots were much lower than in roots. A high As concentration in plant could not directly link to total amount of As in plant. A low As concentration in plant with a big plant size could led to high total amount As

accumulation in plant. Hence, the total As content in 100 15-day old *Arabidopsis* seedlings was also examined (Figure 3.19.c). The results showed that genotypes like Ler-0, although it did not accumulate a high As concentration in plants, because plants grew larger than some other genotypes in the environment presence of arsenite, the total As accumulated was still the highest. The As transfer ability from root to shoot was compared using the concentration of As in plant shoot to against root among all eight hydroponically cultured *A. thaliana* ecotypes (Figure 3.19.d). The ratio between the concentration of As in shoot and root could imply the ability of a plant to transfer the As to the aboveground tissue. In five of the eight genotypes, the concentration of arsenic in the shoots was more than 8% of arsenic in the roots. However, genotype Kas-2 and St-0 had the lowest ration, therefore, their shoot would contain less As content comparing with other tested genotypes under the same root tissue weight.

3.5 Discussion

3.5.1 Analysis of absolute and relative root length change under arsenite stress

Root length was used as the heavy metal stress trait in many studies, especially in seedling stage studes (Chowdhury et al. 2000; Sánchez-Bermejo et al. 2014). It is reasonable since the root is the most developed organ during the seedling stage. Before plants develop the secondary root, the length of the primary root is directly linked to the growth rate (Beadle 2014).

Under arsenite stress, a reduced root length was observed in *B. juncea* and *A. thaliana*. The absolute root length of the plant grown under arsenite environment is direct evidence to indicate which genotype produces longer root under arsenite stress. However, plants have natural variation of root length. When treated with arsenite, the absolute value of root length measurement represents the actual performance of the plant under As stress, which is the consequence of additive effect from both natural growth ability and arsenite tolerance ability (Leão et al., 2014). In order to reduce the impact of the natural variation, data from arsenite treated group and control groups were converted into percentages at the given arsenite concentration. The relative data was the phenotypic change caused by As^{III} in this study. When studying the heavy metal impact, using relative data is more favorable, because the relative data represents the changes directly caused by arsenite stresses, hence genotypes with high relative root length suggests they have better arsenite tolerance than those with low relative root length. The same principle of using relative data was applied when measuring biomass in Chapter 4. The results of the relative root length and the absolute root length in both *B. juncea* and *A. thaliana* were significantly correlated; the R-value was 91.7% and 75.5%, respectively, indicating a very significant effect of arsenite on root length growth.

Data from both *B. juncea* and *A. thaliana* showed that root growth rate decreased with the increasing As concentrations. Root length change was more consistent under higher As concentration. Under lower As concentration, a genotype–environment interaction (G x E) in both *B. juncea* (Figure 3.5) and *A. thaliana* (Figure 3.15) was found, which suggested the plant root behaviour was environmentally dependent under low As stress (El-Soda et al. 2014). One possibility for this could be due to the combined effects of multiple genes and the environment, which affects the expression of genes and trait together (El-Soda et al. 2014). Another possible reason is that plants may mistakenly take up and take on As to phosphate in lower concentration. Phosphate is the essential nutrient in plant growth, and plants may use As in phosphate dependent pathways to grow (Zhao et al. 2010).

Interestingly, after seed germination, very hairy roots were observed in As agar plates on the side where primary roots access the agar surface. Also apical growing points of some roots would resist its geotropism and grew towards the opposite direction, away from the agar plate. Root hairs helps plant to absorb water from air thus to alleviate the stress from As. These observations suggested that measuring root length for As tolerance study may face some problems and interfere with the accuracy of the measurement. The consequence of a big standard deviation can also be reviewed from data analysis of *B. juncea* absolute and relative root length changes (Figure 3.2 and 3.3). The adjusted R^2 in the analysis was 64.30% and 22.2%, respectively, suggested that root length measurement was not accurately linked to seedling arsenite tolerance ability.

3.5.2 *B. juncea* seed and seed oil contained arsenic

When using a hydroponic culture system in As absorption studies in plants, some argued that plant species could be more efficient taking up As from nutrient solutions than under field conditions (Moreno-Jiménez et al. 2010). However, research found that similar or milder effects of As toxicity were obtained in plants grown on soil than in

those grown in nutrient solution. Moreover, the same study also found the shoot-to-root As ratio values were in the same range for both soil and hydroponics (Moreno-Jiménez et al. 2010). These discoveries supported our experimenal design as well as helped us to compare our hydroponic cultivation results with field research.

In this research, a life time 1ppm As(III) (equals to 7.7μM arsenite) exposure was introduced to *B. juncea* since their germination. The setting of the experiment was to simulate the actual environment in Bengal Delta. The average concentration of As in soil in Bangladesh is around 1 to 2ppm (Chowdhury et al. 2000). 4ppm to 14ppm As in seed and 65ppb to 736ppb As in seed oil were observed. Currently, WHO recommended the limit of As in drinking water as 10 μg/litre (ppb) (Curry et al. 2000), is equal to 10ppb. However, the concentration of As in drinking water and in seed oil cannot be directly compared because of the form of As in these two are very different. In drinking water, nearly all As forms are inorganic, which are very toxic. X-ray absorption spectroscopy (XAS) has been used to study the uptake and co-ordination of As in different plant species. One study in *B. juncea* showed a reduction of As(V) to As(III) and the majority of the arsenite remained in the roots as an As(III)-tris-thiolate complex (Castillo-Michel et al. 2007). From the result, it suggested that As is more likely to combine with phytochelatins and convert into non-toxic form in seed oil. Similar findings can also be found in other heavy metal studies, such as Cadmium, where metal ions would combine with PCS protein, which is non-toxic (Heiss et al. 2003). In this research, the seeds and seed oil were analyzed only under 1ppm arsenite stress. Using hydroponic cultivation, another previous research studied the arsenate uptake ability of 14-day old *B. juncea* roots and found 104ppm of As accumulation in the roots, but only 15.3ppm was found in stems and 7.65ppm in shoots, when exposed to 250μM arsenate for 2 days in axenic hydroponic solution (Pickering et al. 2000).

Comparing Figure 3.10. a and b, it indicates that the concentration of As in seeds does not correlate to that in seed oil. Additionally a high As concentration in seed oil

was not directly linked to its oil extraction rate (Table 3.1). However, it is clear that the higher As concentration in seed, the higher As concentration in seed oil. Linking the As tolerance in seedling stage and the As accumulation in seed, it was difficult to find a clear correlation. GP-Jun-1197, for instance, was classified as As-sensitive, because its root elongation ability under As stress was weaker than others (Figure 3.6) and it absorbed very high amounts of As in seed (14ppm). Bar-11 was classified in As tolerance group (Figure 3.6) from the same seedling root length study, and the amount of its absorption of As in its seeds was the second highest (12ppm), only a bit lower than GP-Jun-1197. The oil yield of GP-Jun-1197 and Bar-11 were neither the highest nor the lowest (figure 3.9), suggesting that the high As concentration in oil was not due to high oil extraction rate, so there were other factors interfere with the As transportation from seed to seed oil, and this would require further study.

Literature suggess that tolerance to metals could either be achieved by avoiding the metal stress, or by tolerating it, or both (Patra et al. 2004). No correlation was found between root length change under arsenite stress and As accumulation in seeds. If root length (absolute or relative data) was considered as the parameter of plant As tolerance ability, and As in seed the direct evidence of plant As absorption ability, the two phenotypic evidences suggested that the plant growth ability against the As stress and accumulation of As in the plant were two different phenotype characters and do not link to each other. A complex metabolism is believed behind the phenomenon.

3.5.3 *A. thaliana* exhibited arsenic tolerance in seedling stage and mature plant stage

The examination of root length change under arsenite stress was condacted in Petri dishes, which only allowed plants to grow to seedling stage. The arsenite tolerance ability of mature plants was tested in hydroponic cultivation. Two phenotypic parameters were measured, the leaf number at day 20 in Petri dishes and plant height

at day 45 in the hydroponic system (Figure 3.16). Data of Day 20 represented the development stage before bolting, whereas data of day 45 represented the development stage after bolting. The two phenotypes of 8 *A. thaliana* genotypes had a significant impact from arsenite. A correlation was noticed between the true leaf number and plant height among 8 *A. thaliana* genotypes. This suggests plants exhibit consistent As tolerance ability from different tested environments, as well as at different developmental stages. A study found that mature plant of tomato (*Lycopersicon esculentum*) growth were affected by As and resulted a maximum decrease of 76.8% in leaf fresh weight and maximum reduction of 79.6% in tomato fruit yield when giving a long-term of 10ppm of As^{III} in a nutrient solution (Barrachina et al. 1995). In our experiment, the 4ppm of As^{III} in Hoagland nutrient solution was given to *A. thaliana* plants for 45 days, and significant difference was observed in plant height. Until now, the studies of the effects of arsenic in the different developmental stages of *A. thaliana* plants cannot be found, so it was very difficult to compare our finding with the previous literatures.

3.5.4 Phenotypic changes of *A. thaliana* seeds under arsenite stress

Four As^{III} stress was given to *A. thaliana* plants in hydroponic study to study the effect of As on seeds. Figure 3.17 showed the change of the total seed yield. It seemed that genotypes that had high yield in natural condition would suffer more from arsenite stress. This is probably because more nutrients were required in the growth of high yield genotypes, thus an increased amount of As^{III} was absorbed together with other ions (Zhao et al. 2010). Studies from As and other heavy metals believed that these metals include inhibition of cytoplasmic enzymes and damage to cell structures due to oxidative stress (Chibuike and Obiora 2014). And because As^{III} could be broken down, when concentrations within the plant exceed optimal levels, they adversely affected the plant growth and directly deduced the seed yield (Chibuike and Obiora 2014).

Another finding of this experiment was that the total seed weight pre plant and the

change of total seed weight caused by As revealed difference plant As resistance performances. For instance, the seed weight of As treated group and control group of genotype Col-1 and Col-0 were nearly the same. This result indicated that genotype Col-0 and Col-1 both had a resistance ability to maintain its normal seed production when under As stress. However, when comparing pre plant seed production under arsenite stress, one Col-0 plant can provide 14mg more seeds than a Col-1 plant, which was the1.5 times of a Col-1 plant's total seed production. This suggested that relative data (As+/As-) and absolute data (As+ only) revealed different phenotype again.

3.5.5 Arsenic content in *A. thaliana* root and shoot

ICP-MS analysis was used to measure As content in eight *A. thaliana* genotypes. First, As accumulation was found in both *A. thaliana* root and shoot, which is consistent with previous studies in other plants (Zhu and Rosen 2009). Second, the variation of As accumulation was observed in both root and shoot among genotypes. A variation between difference genotypes under the same experimental condition would suggest a influence at the molecular level led to this phenotypes variation. Third, there is a correlation between concentration of As in root and shoot, implying the more As absorbed by plant roots, the more accumulation of As in shoot. This correlation was significant but not very strong, indicating the existence of different As transport abilities in xylem among *A. thaliana* genotypes.

Combining the data from As content analysis and the data from biomass measurements performed in Chapter 4, a difference in As accumulation between As sensitive types and As tolerant types was discovered. From the change of relative shoot weight under arsenite stress (Figure 4.13), Bur-0 and Kas-2 were the two most As sensitive genotypes, and they were unable to produce seed under arsenite stress from hydroponic cultivation. ICP-MS result (Figure 3.19) suggested that the concentration of As in the Bur-0 root was 1.1 times higher than in Kas-2. In shoot, the

concentration of As in But-0 was 7 times higher than in Kas-2. Bur-0 absorbed more As than Kas-2 when growing under the same As concentration. These two genotypes appear to have the different transport ability (Figure 3.19.d), suggesting Bur-0 has a higher As transport ability in xylem than Kas-2. The As concentration in St-0 roots was the second highest, and the ratio of the As in root and in shoot was the second lowest, suggesting it has low As transport ability in xylem. For the same reason, Nd-1 could also be concluded to have a high As xylem transport ability. Due to author's knowledge, the preliminary analysis of comparing As translocation ability between different *A. thaliana* genotypes or other plant species could not be founded in the literature, therefore, reference of similar studies could not be provide here.

When collecting samples from hydroponic cultivation for ICP-MS analysis, sample weight lose is inevitable when separating tissues from growth media. Consequently, the concentration of As cannot directly link to total As amount in tissues. This shortage could be improved by using agar media. Using the measurement of the average As concentration of 100 15-day old seedlings from eight *A. thaliana* and total plants weight of the 100 seedlings, Figure 3.19.c gives the total As amount of 100 seedling shoots from the eight genotypes. The variations observed from these data are probably caused by different As metabolisms in plants. 100 Ler-0 seedlings accumulated the highest amount of As among the eight tested genotypes. However, Ler-0 did not have the highest As concentration in shoot or in root in hydroponic cultivation. The high As amount in Ler-0 is probably due to its high biomass weight.

3.5.6 Summary

In this chapter, both the arsenite tolerance ability and the absorption level in the crop plant *B. juncea* were first tested. The result confirmed that *B. juncea* showed arsenite tolerance and the possibility to grow *B. juncea* in As polluted land. However, the As content in *B. juncea* seed oil was above the world safety level (10 ppb). Yet, an As concentration variation was discovered among different wild-types of *B. juncea*, which

provides a potential opportunity to breed a *B. juncea* cultivar with a high arsenite tolerance and a low arsenite absorption. The arsenite tolerance behaviour in the model plant *A. thaliana* was examined and variations in both arsenite tolerance and arsenite absorption were also observed among different genotypes. Generally, increasing arsenite concentration in the treatment could cause slower growth and eventually lead to smaller shoot size and shorter root length. Chlorosis was also observed in arsenite treated *A. thaliana* seedlings. Arsenite stress had an impact on *A. thaliana* seed yield and seed vigour. The tested *A. thaliana* genotypes showed an arsenite tolerance variation in seed yield and seed vigour change. *B. juncea* and *A. thaliana* had similar phenotypic changes under arsenite stress, which made it possible to study the arsenite tolerance mechanism in *A. thaliana* and later to apply the finding from *A. thaliana* to the study of *B. juncea*. These findings are novel due to the author's knowledge, which has never be reported in the previous research.

Chapter 4

A. thaliana Arsenite Tolerance Phenotype Study

4.1 Introduction

It has been confirmed that both the oil crop plant *B. juncea* and the model plant *A. thaliana* can tolerate arsenite in the last chapter. Furthermore, a variation of arsenite tolerance was also found among different ecotypes of *B. juncea* and *A. thaliana*. The change of seedling root length in the presence of arsenite was the main parameter used for measuring arsenite tolerance. However, other phenotypic character changes of plant seedlings were also observed when given arsenite. In this chapter, the research was to quantify the arsenite tolerance of *A. thaliana* by analyzing the change of shoot biomass accumulation, chlorophyll fluorescence as well as anthocyanins fluorescence in the presence of arsenite. Whether anthocyanins were involved in As detoxification was also studied. Results from comparing arsenite tolerance ability of various *A. thaliana* genotypes could then be used to select tolerance and susceptible genotype, which was the precondition to generate mapping population for genetic mapping of arsenite tolerance loci.

The knowledge from the phenotype analysis mentioned above would be essential in arsenite tolerance index selection. The optimal index can increase the accuracy of genetic mapping and provide a higher chance to find a specific gene or locus for arsenite tolerance. In order to select suitable genotypes for generating RIL population and to perform GWAS analysis, the measurements of arsenite tolerance of various *A. thaliana* genotypes were taken.

4.1.1 Plant pigments change under arsenic stress

Plants accumulate anthocyanins under As stress, as well as many other abiotic stress. One study using *Azolla caroliniana* to investigate the impact of 20 μM As on plants, observed a 30% decrease in biomass production and an increase in the concentration of anthocyanin (Rofkar et al. 2014). Another study on *Lemna gibba* showed the anthocyanin content increased constantly along with increasing As concentration

(Leão et al. 2014). Anthocyanins are a subgroup of flavonoids, which are responsible for the display of pigments associated with leaf senescence (Montefiori et al. 2011). Juszczuk et al. (2004) studied bean plants under As stress, and claimed that anthocyanins played an important protective role in plants, as they are produced in the response to oxidative stress. However, the study did not investigate details of the actual defence metabolism, nor provided any defence metabolism hypothesis. And studies of function of anthocyanins in plants under abiotic stress cannot be found in the literature.

AT5G13930 is a gene in *A. thaliana* encoding chalcone synthase (CHS). Therefore it is required for the accumulation of purple anthocyanins in leaves and stems (Figure 4.1). The *tt4* mutant is deficient in anthocyanins accumulation, as it is a null mutation of the chalcone synthase (CHS) gene (Peer et al. 2001). In this chapter, the flavonoid-deficient *tt4* mutant is used to investigate the consequences of the lack of anthocyanins under arsenite stress.

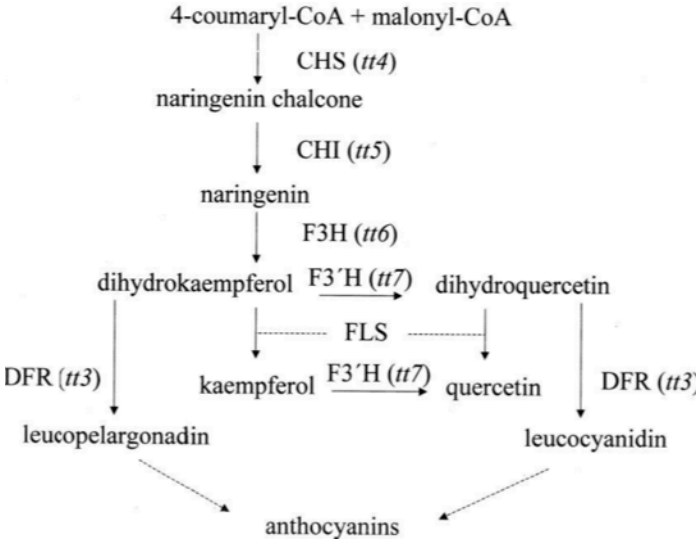

Figure 4.1. Schematic diagram of the flavonoid biosynthetic pathway (Adapted from Peer et al. 2001).

From the study of the As effect on the content of other pigments, Leão et al. (2014) also found that the concentrations of chlorophyll a, chlorophyll b, and total chlorophyll were significantly reduced in plants exposed to As, but that the concentration of

carotenoids was unchanged. Etiolation in the shoot is often observed under As stress, because As can lead to disruption in energy flows in cells after it is transported into the cytoplasm (Päivöke and Simola 2001; Singh et al. 2006; Rahman et al. 2007 and Stoeva et al. 2003). However, correlation between the change of chlorophyll and the change of As stress has been rarely quantified in previous research work.

4.1.2 Review of two arsenic related genes

The plant transportation and detoxification mechanism of As were reviewed in Chapter 1. These pathways are not exclusive to As but probably include many heavy metals, such as chromium and mercury. Furthermore, the chemical properties of arsenate enable its transport into the cytoplasm with phosphate (details of the process have been given in Chapter 1, Section 1.5.1), whereas arsenite, chromium and mercury, are taken up via aquaporin channels (see Section 1.5.2 for more details). One general metabolism to detoxify As is through chelation (see Section 1.5.5.1) and sequestration of As(III)-PC_2 in vacuolaes (Section 1.5.5.2). AtPCs1 is a phytochelatin synthase in *A. thaliana* (Vatamaniuk et al. 1999). The expression level change of this gene under arsenite stress and control was tested in this chapter. An increased expression level of this gene under arsenite stress should be observed if it helps plant to tolerant arsenite.

Except gene AtPCs1, which is related to plant As detoxification, recently, another gene was also discovered to help plant to tolerant arsenate. Using genetic mapping, a study discovered a gene (At2g21045, ATQ1) that encoding arsenate reductase in *A. thaliana* (Sánchez-Bermejo et al. 2014). The arsenate reductase helps plants to convert arsenate to arsenite, so it can increase plant As detoxication ability (Sánchez-Bermejo et al. 2014). In the research of Sánchez-Bermejo et al. (2014), *A. thaliana* seedlings were grown on Johnson medium in phosphate starvation condition (no phosphate) for 7 days and transferred to 15 μM As(V) medium for 1 day. Plants were then grown in 1mM phosphate plates in a vertical position for another day. The

percentage of plants that showed Root Growth Recovery (%RGR) was quantified and employed for QTL mapping under different concentrations of As^{V} (e.g. Figure 4.2) and As^{III}. In the research of Sánchez-Bermejo et al. (2014), it can be confirmed that Col-0 was a tolerant genotype under both As^{V} and As^{III} stress. When measuring root length change under As^{III} stress using RGR assay, it appeared Kas-1 acted as similar as Col-0 under As^{III} stress and they had developed similar root length under all tested AsIII concentration (Figure 4.3.b). In this chapter, the arsenite tolerance of Kas-1 was examined through examining the shoot weight change.

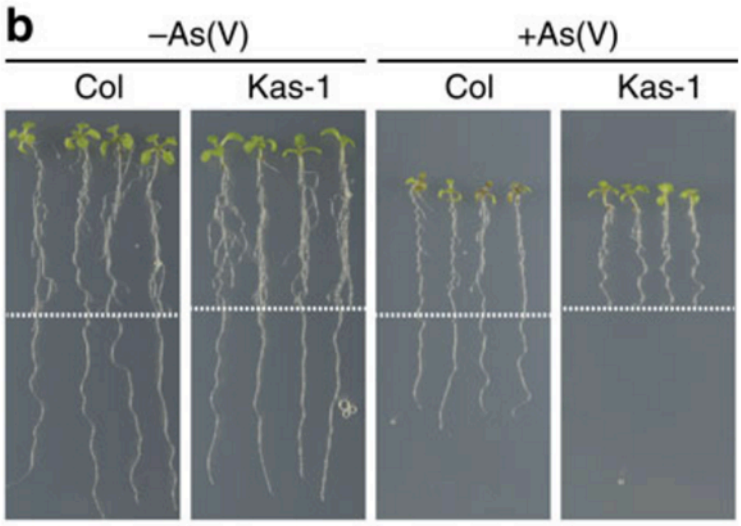

Figure 4.2. As(V) tolerance phenotype of Col *gl-1* (Col) and Kashmir (Kas-1) plants using the RGR assay (Sánchez-Bermejo et al. 2014).

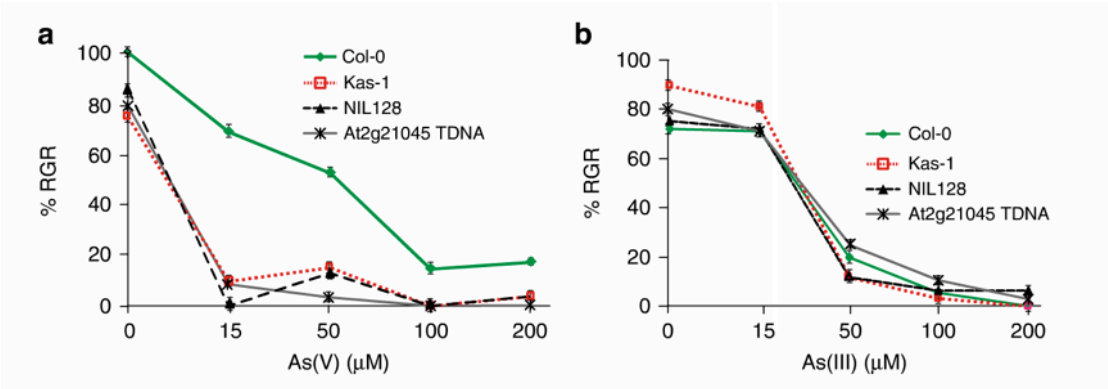

Figure 4.3 a. As(V) and b. As(III) tolerance analysis of Col-0, Kas-1, NIL128 and the At2g21045 T-DNA mutant. Percentage of RGR (%RGR) was determined at various As(V) and As(III) concentrations (Sánchez-Bermejo et al. 2014).

Two mapping populations were used in this research (Sánchez-Bermejo et al. 2014), the Col gl-1 x Kas-1 and Col-0 x Tsu-0 populations. Kas-1 and Tsu-0 were the As(V)-sensitive parents. By QTL mapping, a gene named At2g21045 (ATQ1) was found that encodes arsenate reductase, which is an essential component for natural plant variation in As(V) tolerance (Sánchez-Bermejo et al. 2014). It was necessary to test the phenotypic change when plants lack this arsenate reductase gene under arsenite stress. In this chapter, the plant arsenite tolerance ability in the absence of this gene was also tested through comparing the arsenite tolerance of T-DNA insertional mutagenesis *A. thaliana* plant and its background plant Col-0 genotype. Similar arsenite tolerance level in the T-DNA insertion plant and Col-0 genotype would indicate the existence of another mechanism for arsenite tolerance in plants.

The questions regarding the phenotypic change caused by arsenite in *A. thaliana*, the function of anthocyanins in *A. thaliana* plant and the function of genes AtPCS1 and ATQ1 in presence of arsenite were addressed in this chapter. In addition, the variation of arsenite tolerance was examined among genotypes of *A. thaliana*.

4.2 Materials and Methods

4.2.1 Materials

All *A. thaliana* genotypes used in this chapter were listed in Section 2.1.1.2.

4.2.2 Main methods

The major method used in this chapter were:

Phenotype change observation assay;

Rosette measurement of *A. thaliana*;

Post-germination assay of *A. thaliana*;

Chlorophyll content of *A. thaliana* seeding;

Anthocyanin content of *A. thaliana* seedling;

RT-PCR analysis;

Agarose gel electrophoresis;

Analysis of T-DNA insertion plant line.

The detailed procedures of these experiments were given in Chapter 2.

4.3 Results

4.3.1 Changes of phenotypic characters under arsenite stress

To observe the effect of arsenite on plants, *A. thaliana* were cultured in seedling growth media in 300ml volume (68mm diameter x 100mm height) tissue culture pots to let seeds germinate and grow (for the method, refer to Section 2.2.12). The capacity of the pot allowed a long-term growth of plants with long-term arsenite stress. 14 days after germination, a clear difference between the control group and As stressed group was observed. Data of two *A. thaliana* genotypes Col-0 and Tsu-0 in control and 4ppm, 6ppm arsenite treatment can be seen in Figure 4.4.

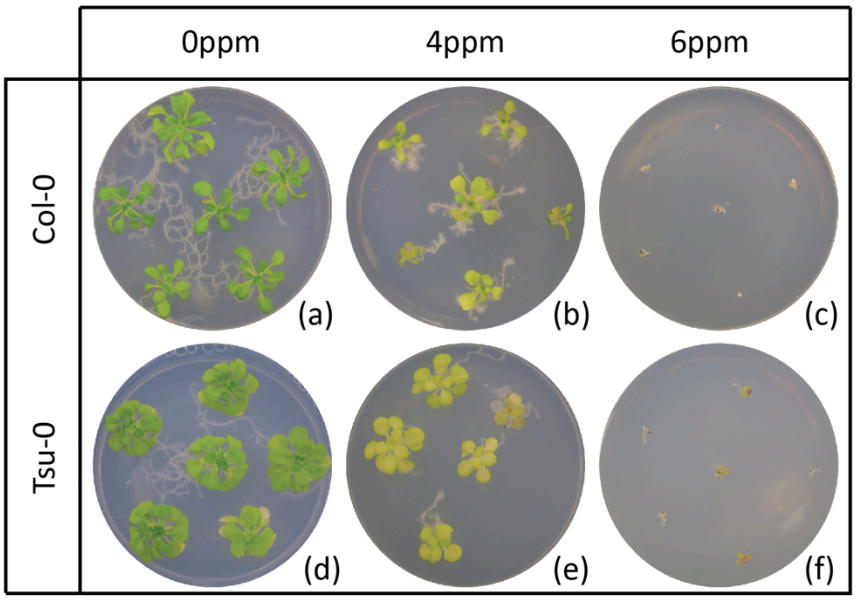

Figure 4.4. Representative images of two 14-days old *A. thaliana* genotypes Col-0 and Tsu-0 in control conditions and under arsenite treatment. a. Col-0 in 0ppm arsenite; b Col-0 in 4ppm arsenite; c. Col-0 in 6ppm arsenite; d. Tsu-0 in 0ppm arsenite; e: Tsu-0 in 4ppm arsenite; f: Tsu-0 in 6ppm arsenite.

In the control groups, all seedlings of each genotype grew healthily and consistently. In the 4ppm arsenite treated group, seedlings of both genotypes were smaller, with shorter rosettes. Etiolation was observed at this stage. The shoot size and germination rate were not consistent. In the 6ppm arsenite treated group, seeds

barely germinated and even if they did, they grew very slowly. *A. thaliana* seedlings were left to grow for one month. During this period, all seedlings from the 6ppm arsenite treated group could only develop one pair or two pairs of true leaves; meanwhile, the seedlings also became whitened and eventually stopped growing.

If seeds were germinated in arsenite free media and then transferred into arsenite media, seedlings could grow under a 6ppm arsenite environment. An accumulation of purple pigment was observed (Figure 4.5.c) on the bottom of the seedling leaves, which suggested an accumulation of anthocyanin in the leaves.

Figure 4.5. Representative images of 14-day old seedlings of *A. thaliana* seedlings. Among all 23 tested genotypes, Kas-2 genotype shows the strongest purple colour, whereas Col-0 shows the least. a. Kas-2, 0ppm arsenite and b. Bur-0, 0ppm arsenite; c. Kas-2, 6ppm arsenite and d. Bur-0, 6ppm arsenite. Arsenite source: sodium arsenite. Scale bar is 1mm.

4.3.2 Will larger seeds do better under arsenite stress?

A deviation of seedlings size from the same genotype was observed in response to

As in chapter 3. The deviation could be due to the fact that arsenite can influence the germination speed and also reduce the rosette diameter. To study the correlation between the seed mass and seed vigour change under arsenite stress, four *Arabidopsis* genotypes with different seed biomass were selected from 22 genotypes (Observation from data in Section 3.4.3).-Among all genotypes, Bur-0 and Kas-2 had the largest seeds (Figure 4.6.a & b), while Ler-0 and Mt-0 had the smallest seeds (Figure 4.6.c & d).

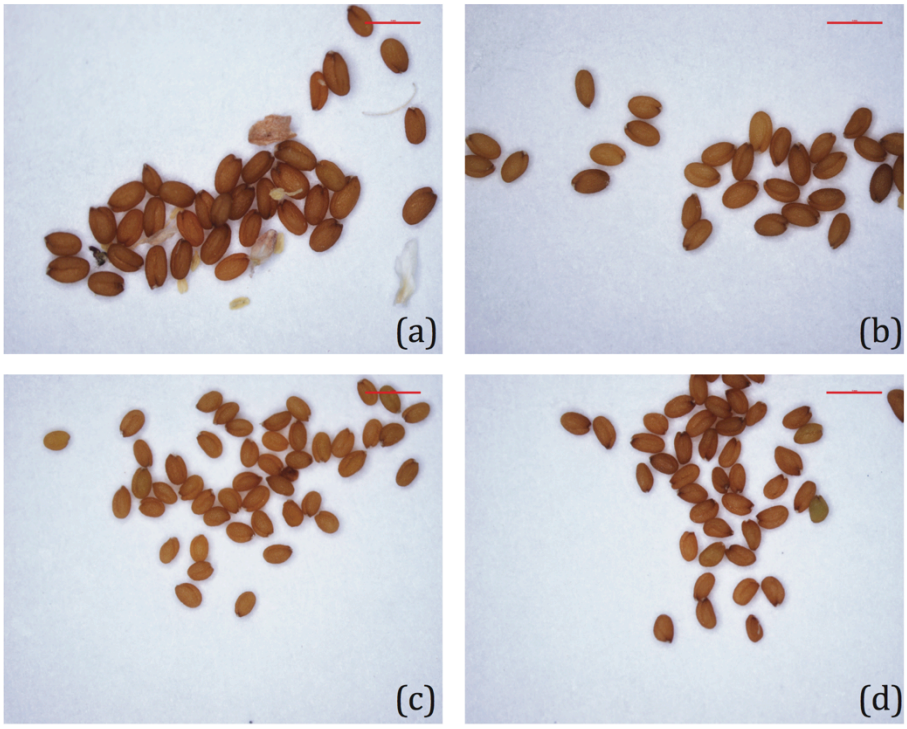

Figure 4.6. Images of mature seeds from four different genotypes. a. Bur-0 genotype; b. Kas-2 genotype; c. Ler-0 genotype and d. Mt-0 genotype. Scale bar: 1mm.

Seeds from Bur-0, Kas-2, Ler-0 and Mt-0 were artificially modified (for the method, refer to Section 2.2.2.2) in order to get larger seeds. Both the weight and size of normal-seed and larger-seed were examined to make sure that the modification was successful. The measurement of weight and size of normal and larger seeds are summarized in Figure 4.7.

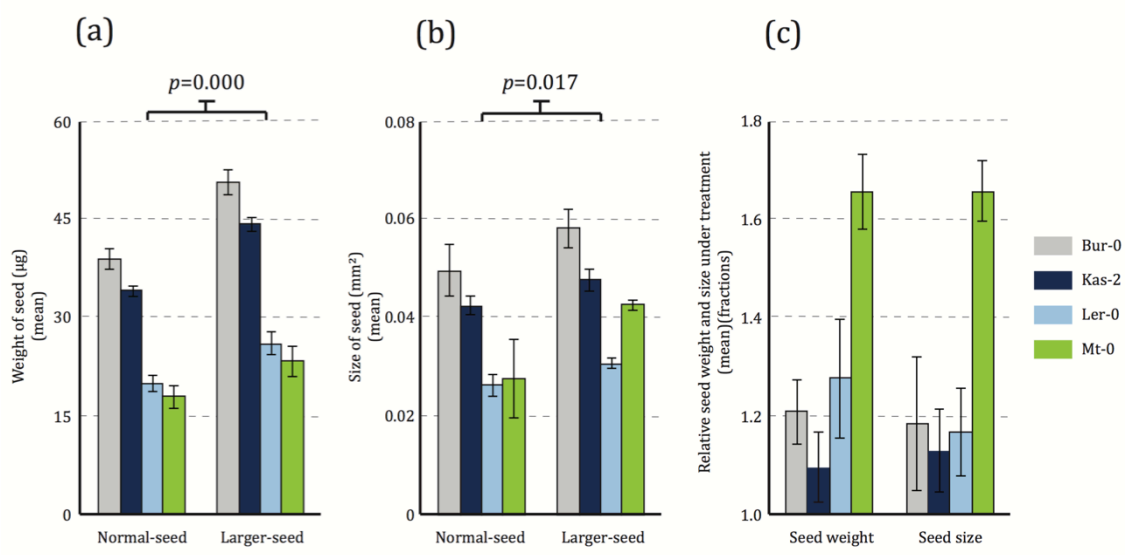

Figure 4.7. Comparison of normal-seeds and larger-seeds of 4 *A. thaliana* genotypes. a. The mean weight of seed; b. The mean size of seed; c. The difference between normal-seeds and larger-seeds in seed weight and size. Data are means of five replications (±SD).

It has been verified that the seed modification (Section 2.2.2) can significantly increase seed weight and size. Comparing four genotypes, a significant difference in seed weight and size between normal-seed group and larger-seeding groups was seen. In normal seed group, *P*-values of weight analysis and size analysis were both 0.000. In larger-seed groups, the difference between Bur-0, Kas-2 and Ler-0, Mt-0 was still significant (Independent sample T-test, seed weight *P*-value = 0.000; seed size *P*-value = 0.000). Comparing the difference in seed weight and size change (Figure 4.4.c), it was very clear that weight and size of Mt-0 genotype had the largest change in seed weight and size.

The germination rate of the seed of both normal and larger group was measured under controled conditions (arsenite free) and arsenite conditions (4ppm sodium arsenite) at 48 and 72 hours after seeds being laid (Figure 4.8).

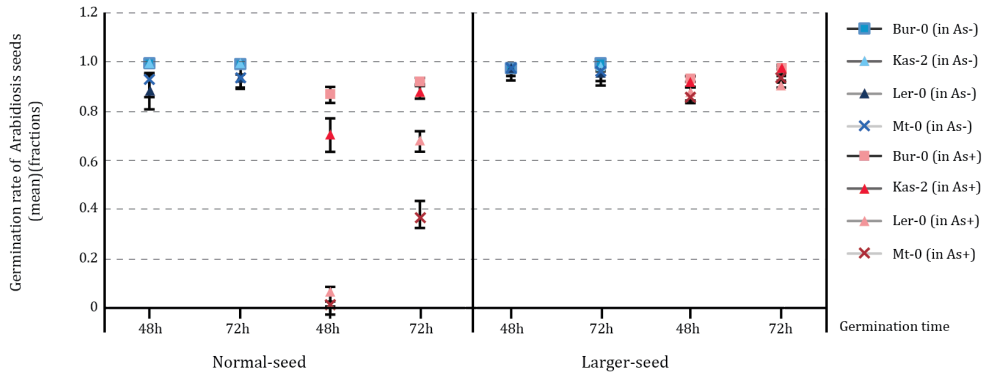

Figure 4.8. The germination rate (100 seeds) of 4 *A. thaliana* genotypes. Data was collected at 48 hours and 72 hours. Seeds germinated in 4ppm sodium As group and control group are both presented. Data are means of five replications (±SD).

In the arsenite free (control) environment, the germination rates of normal-seed (blue symbols) from Ler-0 and Mt-0 were lower than Kas-2 and Bur-0 at the 48h, but not significantly different. This difference became smaller at the 72h. In the larger-seed group, a difference of germination rate of all four genotypes between the 48h and the 72h was rarely observed.

Analyzing the germination rate of normal-seed in the arsenite environment, a significant difference was found (red symbols) between 48h and 72h (Independent sample T-test, *P*-value = 0.024), especially for Ler-0 and Mt-0, which had no seed germinated at the 48h time point. However, there was no significant difference in in larger-seed germination rate between 48h and 72h. After 72 hours, nearly all seeds from larger-seed groups germinated.

Comparing germination rate of the normal-seed and larger-seed in arsenite treated group (red symbols) and the arsenite free (control, blue symbols) group, the former were always lower than latter. Analyzing normal-seed germination, a significant increase of germination rate was found between 48h and 72h (Independent sample T-test, *P*-value = 0.000). This suggested that arsenite caused germination delay.

Focusing on Ler-0 and Kas-2, increasing seed biomass could significantly increase the germination rate at 48h (Independent sample T-test, *P*-value = 0.000), suggesting the small seeds had longer germination delay than larger seeds.

4.3.3. Arsenite tolerance in post-germination stage

This experiment compared the As^{III} tolerant ability between seedlings that germinate from the seeds with different size. Theoretically, seeds with naturally large size carry more nutrients than seeds with small size and develop bigger seedlings after germination. Whether the size of the seed interferes the seedling size under As stress was analysed in this section.

The post-germination stage was defined as the growth between 7-days and 14-days in this research. The seedling weight of four genotypes, including normal seeds and larger seeds which were analyzed in Section 4.3.2, was measured to study the post-germination stage in both the 4ppm arsenite treated group and the control. (Figure 4.9)

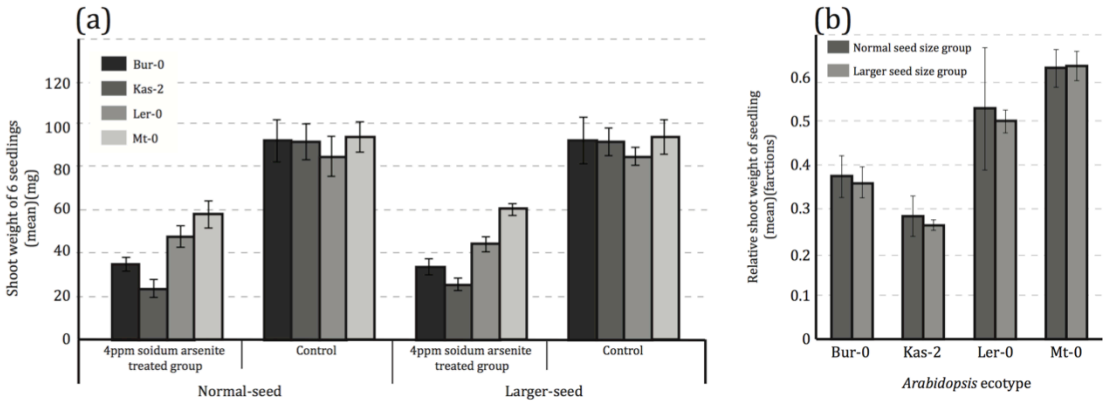

Figure 4.9. a. The mean seeding shoots weight of 4 *A. thaliana* genotypes from post-germination assay in 4ppm sodium arsenite and control. (4 replicates) b. The change of shoot weight caused by As^{III} (relative change) of 4 *A. thaliana* genotypes under 4ppm of sodium arsenite stress. Data are means of five replications (±SD).

In Figure 4.9, Bur-0 and Kas-2 represented the larger seed group, because they had large seed in natural environment as well as after artificially modification (for the method, refer to Section 2.2.2.2). During germination, Bur-0 and Kas-2 had faster germination speed as well as high germination rate than Ler-0 and Mt-0 under As stress (Figure 4.8). However, the analysis of seedlings growth against As showed the different result. At the post-germination stage, the absolute shoot weights (Figure 4.9.a) and relative shoot weights (Figure 4.9.b) of Bur-0 and Kas-2 were much lower than Ler-0 and Mt-0. It appeared that the advantage from large seed size in germination stage did not help Bur-0 and Kas-0 to develop to big seedlings. No significant difference in seedling weight could be found between the shoot of seedlings grew from normal seeds and artificially modified seeds (Paired samples T-test, *P*-value=0.071). This result suggested that seed size was not a factor to correlated to plant seedling As resistance ability.

4.3.4 Anthocyanin content changed under arsenite stress

To study the correlation between seedling shoot weight change and anthocyanin content change under arsenite stress, three physiological indexes were measured in 6ppm sodium arsenite and under control conditions. The anthocyanin content was extracted from all the arsenite treated seedlings. In the control (arsenite free) group, there was no anthocyanin detected in the seedlings. All data are presented in Figure 4.10.

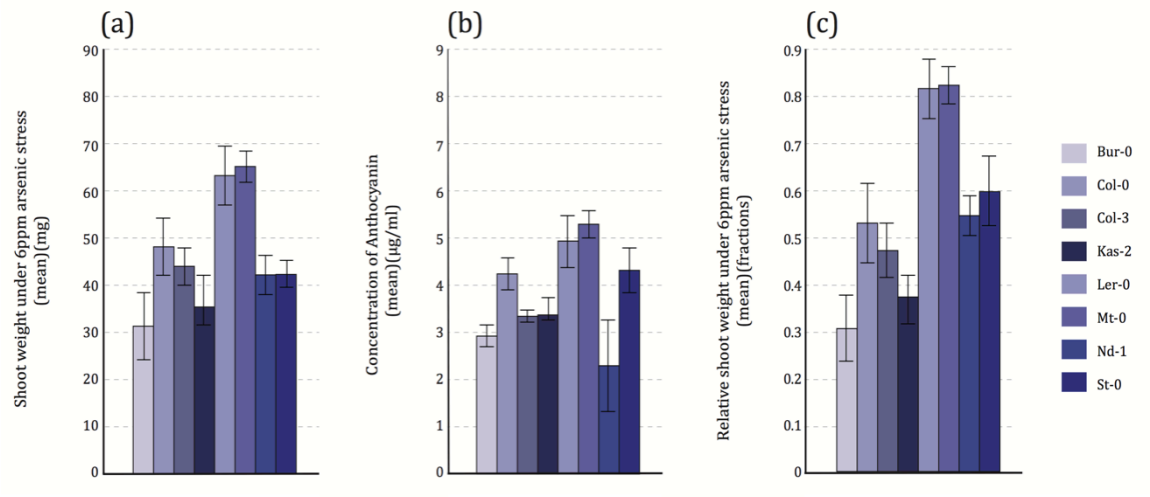

Figure 4.10. a. The absolute weight of 6 shoots of 14-day old seedlings of 8 *A. thaliana* genotypes under 6ppm sodium arsenite stress; b. The anthocyanin content extracted from 14-day old seedlings, which was given the 6ppm sodium arsenite stress; c. The change of shoot weight caused by As^{III} (relative change) of 14-day old seedlings under the 6ppm sodium arsenite stress. Data are means of four replications (±SD).

The results suggest a significant difference in shoot weight change among 8 *A. thaliana* genotypes, which was caused by 6ppm sodium arsenite (One-way ANOVA, *F*-value = 15.500, *P*-value = 0.000). The total anthocyanin amount extracted from arsenite treated seedlings among all *A. thaliana* genotypes was significantly different (One-way ANOVA, *F*-value = 13.501, *P*-value = 0.000). There was also a significant difference in relative shoot weight change under arsenite stress (One-way ANOVA, *F*-value = 27.187, *P*-value = 0.000). The data of the three groups are significantly correlated and are presented in table 4.1.

Table 4.1 The Pearson correlation among absolute shoot weigh, anthocyanin content and change of shoot weight caused by As^{III} (relative shoot weight) under 6ppm sodium arsenite treatment.

		Seedling weight	**Anthocyanin content**	**Relative shoot weight**
Seeding weight	Pearson correlation		0.750**	0.940**
	P-value		0.000	0.000
Anthocyanin content	Pearson correlation			0.0731**
	P-value			0.000

**. Correlation is significant at 0.01 level.

Under arsenite stress, *A. thaliana* produced anthocyanin. Some genotypes accumulated anthocyanin more than others. The anthocyanin content of a plant was correlated to the plant's shoot weight. However the function of anthocyanin in plants is not clear. Comparing *A. thaliana* Col-0 genotype and *a* T-DNA insert mutant *tt4* (Col-0 background) which could not produce any anthocyanin, the change of shoot weight between Col-0 and *tt4* mutant was measured in Figure 4.11.

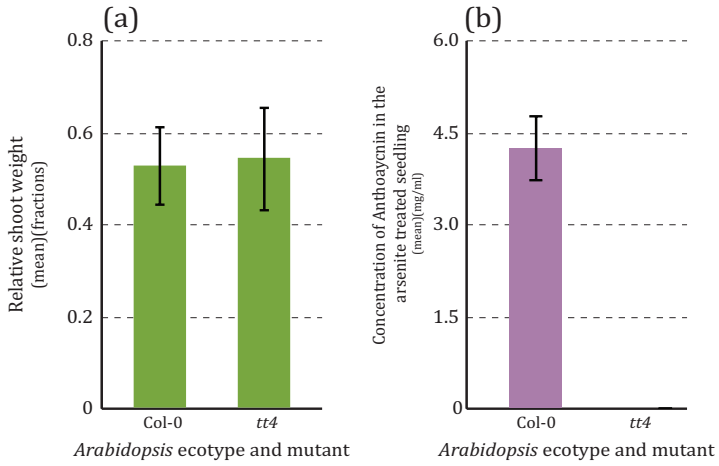

Figure 4.11. a. The change of shoot weight caused by As^{III} (relative change) of Col-0 genotype and *tt4* mutant in 6ppm sodium arsenite; b. The anthocyanin content from 6ppm sodium arsenite treated seedling. Data are means of four replications (±SD).

There was no significant difference in the change of relative shoot weight between Col-0 and *tt4* (Independent sample T-test, *P*-value =0.187). Because the *tt4* mutant could not produce any anthocyanin, but had the similar shoot weight as the Col-0 genotype, it suggests that anthocyanin was not involved in plant arsenite detoxification, but is a by-product under arsenite stress.

4.3.5 A novel way to measure arsenite stress

The results from Section 4.3.4 suggested that measuring seedling shoot weight change under arsenite stress was one way to measure arsenite tolerance. Thus, a post-germination assay was developed to consistently measure arsenite tolerance (procedure presented in Section 2.2.14). The difference of shoots size and green pigment between tested genotypes were both observed (e.g. Figure 4.12). By the measurement of seedling shoot weight and chlorophyll content, the two physiological parameters were quantified.

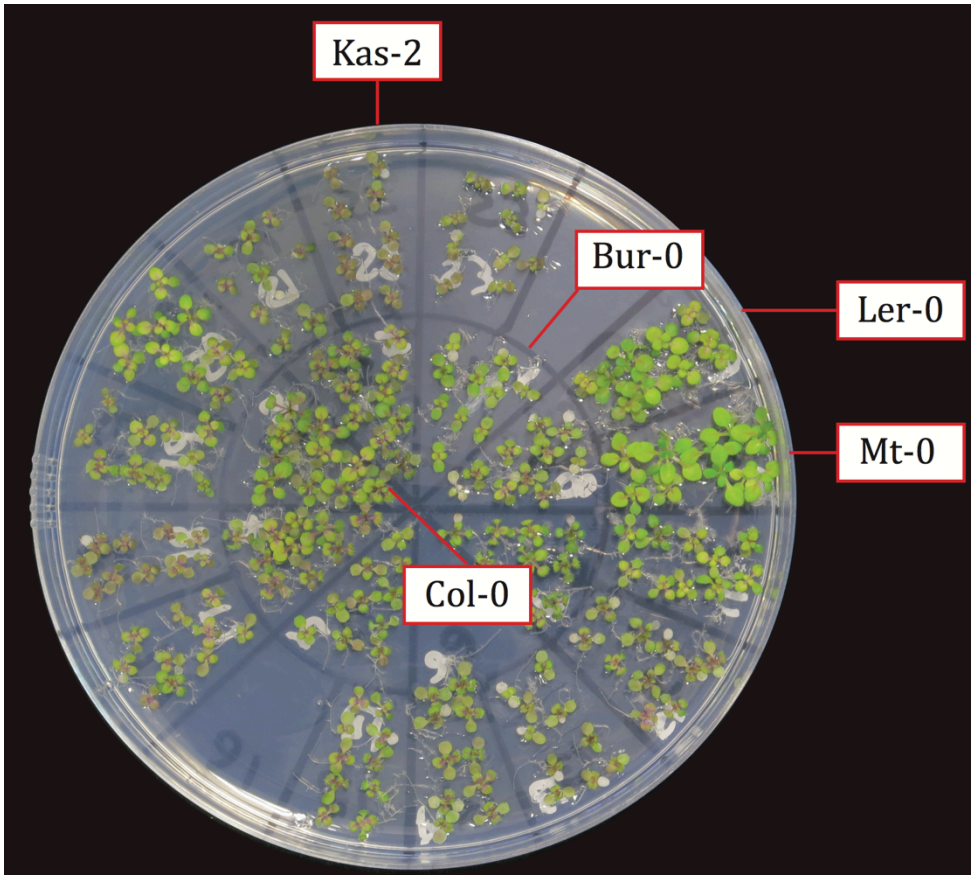

Figure 4.12. A representative image of a post-germination assay of 21 *A. thaliana* genotypes growing under 6ppm sodium arsenite stress.

The analysis of chlorophyll including chlorophyll a, chlorophyll b， total chlorophyll and their content in seedling is presented in Figure 4.13.a. The relative shoot weight change is compared in Figure 4.13.b.

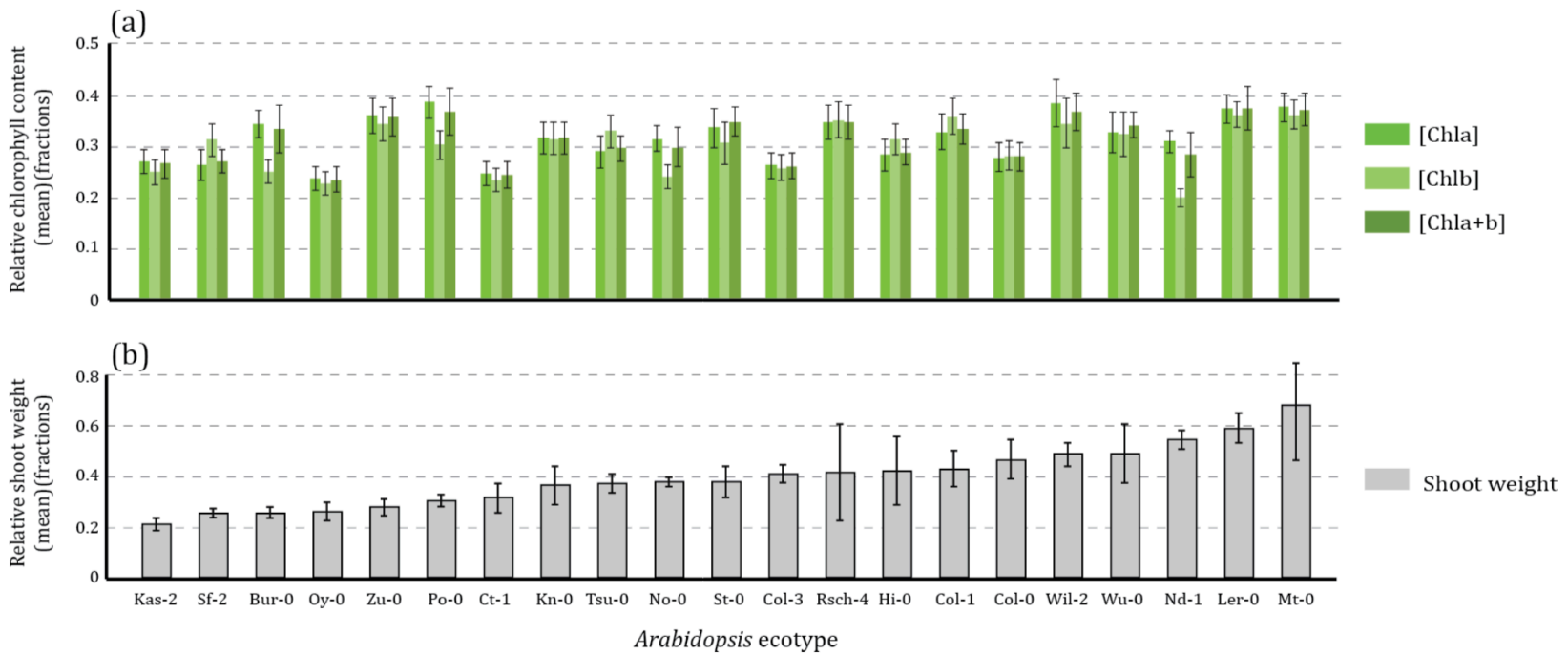

Figure 4.13. a. The average relative chlorophyll content change by arsenite (the chlorophyll content of arsenite treated group: the chlorophyll content of the control group). Measurement was collected from 14-day old seedlings of 21 *A. thaliana* genotype under 6ppm sodium arsenite stress and the control. Data are means of four replicates (±SD); b. The average relative seedling weight change by arsenite (the seedling weight of arsenite treated group: the seedling weight of the control group). Measurement was collected from 14-day old seedlings of 21 *A. thaliana* genotype under 6ppm sodium arsenite stress and the control. Data are means of four replicates (±SD).

A significant difference was observed in the change of relative shoot weight (One-way ANOVA, *F*-value=6.105, *P*-value=0.000). A significant difference was also found in natural shoot weight (One-way ANOVA, *F*-value=6.133, *P*-value=0.000) and absolute shoot weight under arsenite stress (One-way ANOVA, *F*-value=6.684, *P*-value=0.000) among 21 genotypes. Analyzing total chlorophyll content, a significant difference was found among 21 genotypes. (One-way ANOVA, *F*-value=4.658, *P*-value=0.000)

A cluster analysis was carried out using mean relative shoot weight change of all 21 *A. thaliana* genotypes under 6ppm arsenite stress in order to classify the arsenite tolerance ecotypes and the arsenite sensitive ecotypes and divide those genotypes into 3 groups (Figure 4.14). The group 1 was the arsenite susceptible group; the group 2 had middle level arsenite tolerance; group 3 was the arsenite tolerant group.

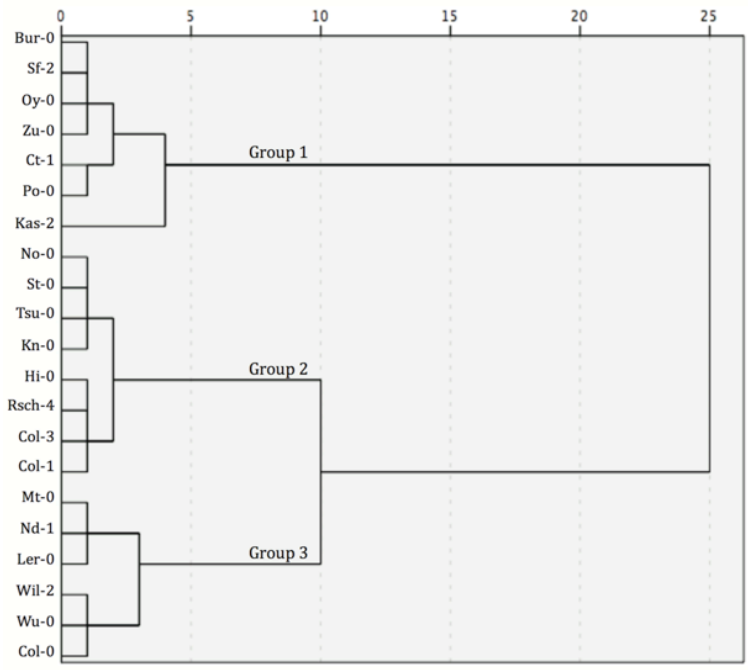

Figure 4.14. The cluster analysis of change of shoot weight by 6ppm sodium arsenite of 21 *Arabidopsis* genotypes. Group 1 has the lowest As tolerance; group 2 has the medium As tolerance; group 3 has the height As tolerance.

Measurement at one concentration only reviews the response at that concentration. Different concentrations were used to determine a response curve for measuring

plant arsenite tolerance and compared the shoot weight of the distribution to obtain a lethal concentration 50 (LC50) where the growth is 50% reduced.

Two genotypes Bur-0 and Kas-2 from group 1 (arsenite sensitive), two genotypes Col-3 and St-0 from group 2 (middle) and four genotypes, Mt-0, Nd-1, Col-0 and Ler-0, from group 3 (arsenite tolerance) were selected to analyze their LC50. Firstly, the mean shoot weights of eight genotypes of *A. thaliana* under a range of sodium arsenite concentrations were calculated (Figure 4.15). Then the trend line formulas of each genotype were calculated based on Figure 4.15 (Table 4.2). R^2 of each trend line formula is also given in Table 4.2. And a half lethal concentration was calculated using the trend line formula and is presented in Figure 4.16.

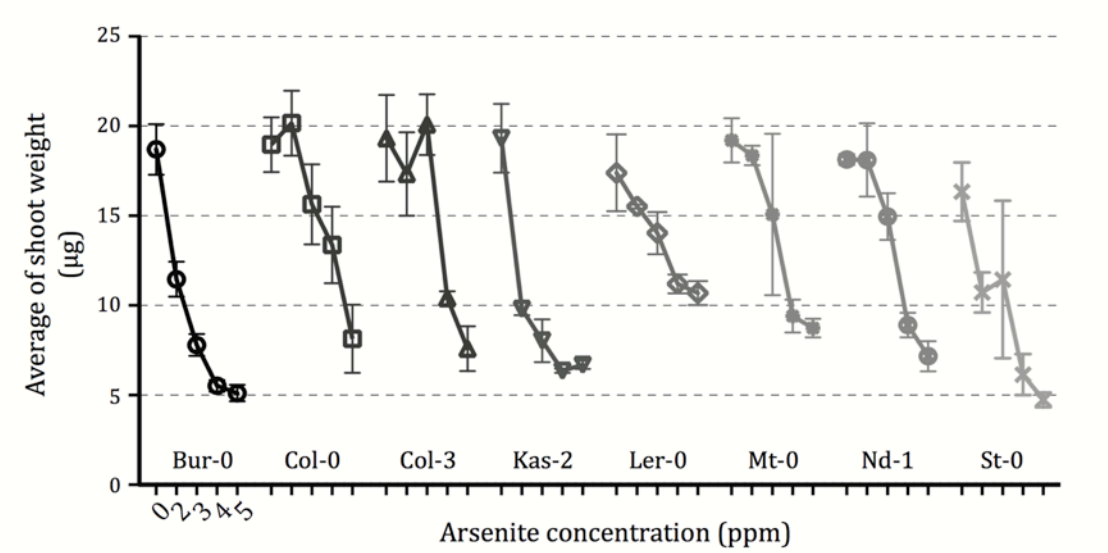

Figure 4.15. The average shoot weight under 2ppm, 3ppm, 4ppm and 5ppm of sodium arsenite stress and 0ppm (arsenite free) control group. Data are means of six replications (±SD).

Table 4.2 The trend line formula and R^2 calculated from Figure 4.15.

Genotype	Trend line formula	R2
Bur-0	y= -3.3116x +19.654	0.86853
Col-0	y= -2.845x +23.793	0.8775
Col-3	y= -3.0392x +24.068	0.7334
Kas-2	y= -2.8721x +18.661	0.71651
Ler-0	y= -1.7723x +19.087	0.97106
Mt-0	y= -2.9887x +23.127	0.93123
Nd-1	y= -3.1143x +22.8	0.9144
St-0	y= -2.7796x +18.218	0.90513

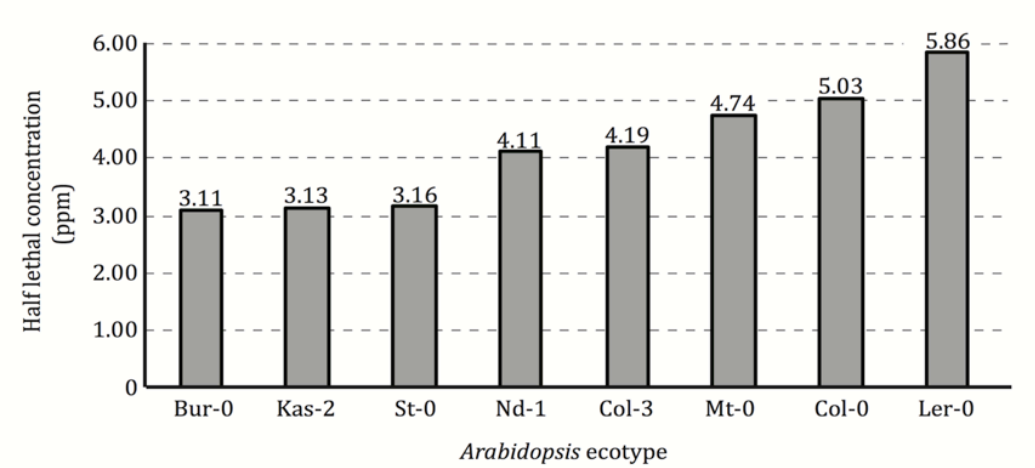

Figure 4.16. The half lethal concentrations (LC50) of 8 *A. thaliana* genotypes.

Analysis of LC50 was one way to interpret shoot weight and a definite comparison of arsenite tolerance of the eight genotypes of *A. thaliana* can be made. From the LC50 data, Ler-0 had the highest arsenite tolerance, Col-0 ranked the second highest; Bur-0 had the lowest arsenite tolerance, and Kas-2 ranked the second lowest (Figure 4.16).

4.3.6. Change of the rosette under arsenite stress

It was observed from Figure 4.4 that the rosette diameter of 14-day old seedlings showed a change under arsenite stress. The same eight *A. thaliana* genotypes as in table 4.2 were studied for rosette diameter change. The concentration of 4ppm of sodium arsenite was supplied to the seedling growth medium. The average rosette diameter of control and arsenite group is shown in Figure 4.17.a. Relative rosette change was compared in a Boxplot analysis (Figure 4.17.b).

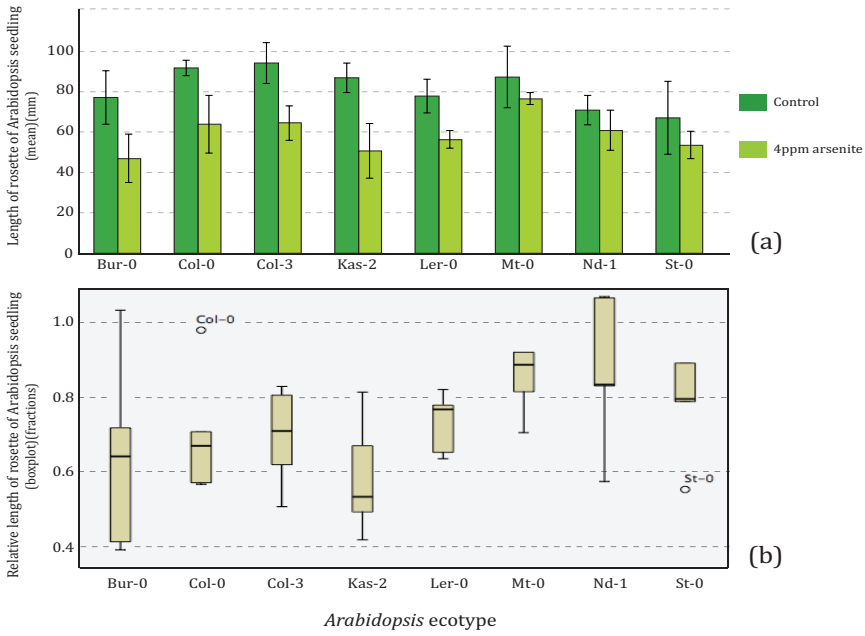

Figure 4.17. a. The mean diameter of rosette (mm) of 14 days old seedling of 8 *Arabidopsis* ecotypes in control and 4ppm sodium arsenite treatment. b. The boxplot of change of rosette diameter by As^{III} of 8 *Arabidopsis* ecotypes. Data are means of four replications (±SD).

By comparing the arsenite treated group and the control group, a significant difference was found in the absolute rosette diameter in both natural variations (One-way ANOVA, *F*-value=3.752, *P*-value = 0.004), as well as the arsenite treated group (One-way ANOVA, *F*-value 4.544, *P*-value = 0.001). When analyzing the relative rosette diameter change, no significant difference was found in the change of

relative rosette diameter (One-way ANOVA, *F*-value=1.945, *P*-value = 0.095).

4.3.7 Investigation of AtPCs1 gene under arsenite stress

Phytochelatin has been reported to play a major role in As detoxification. AtPCs1 is the gene encoding phytochelatin synthase (PCS) protein. To study the expression change of AtPCs1 in the presence of arsenite, a semi-quantitative PCR was performed using housekeeping gene GAPC as the internal control. The cDNA of Bur-0, Kas-2, Ler-0, and Mt-0 genotypes was synthesized from its RNA, which was extracted from seedling shoots treated with 6ppm sodium arsenite and under control conditions. In theory, the brightness of the band is linked to amplified cDNA concentration.

After 20 PCR amplification cycles, there was almost no difference in the brightness of AtPCs1 bands among four genotypes (Figure 4.18). The same conclusion could be drawn from the result of 30 amplification cycles (Figure 4.19). The results indicated that the concentration of amplified AtPCs1 cDNA had no significant difference. Thus, the expression of AtPCs1 was not up-regulated under 6ppm sodium arsenite stress.

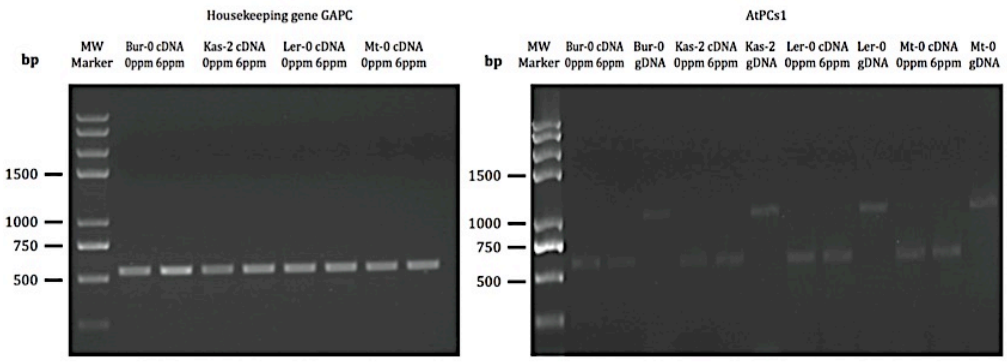

Figure 4.18. GAPC (left) and AtPCs1 (right) amplification produced by PCR after 20 cycles of cDNA synthesized from 4 different *A. thaliana* genotypes in both 6ppm sodium arsenite treated and control groups. AtPCs1 amplification generated siliques by PCR of genomic DNA were also used as the positive control. PCR products were analyzed by electrophoresis through a 1.2% (wt/vol) agarose gel. Electrophoresis was conducted during 50 min at 100 V. The molecular marker used was the 1 kb DNA ladder marker.

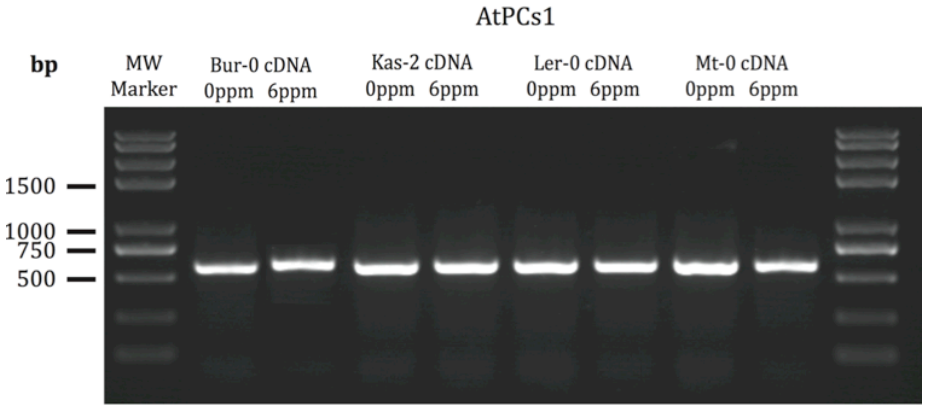

Figure 4.19. AtPCs1 amplification produced by PCR after 30 cycles of cDNA synthesized from 4 different *A. thaliana* genotypes in both 6ppm sodium arsenite treated and control groups. PCR products were analyzed by electrophoresis through a 1.2% (wt/vol) agarose gel. Electrophoresis was conducted during 30 min at 100 V. The molecular marker used was the 1 kb DNA ladder marker.

4.3.8 Arsenite tolerance of *Arabidopsis* T-DNA insertion line At2g20145

Previously, one research has found a gene encoding an arsenate reductase in *Arabidopsis thaliana* that helped plants to resist arsenate stress. To study the function

of this gene (At2G21045) under arsenite stress, the relative shoot weight change under 4ppm arsenite stress was scanned, which was approximately equal to 30μM As(III). *A. thaliana* wild types *A. thaliana* Col-0, Col-5(gl1), Kas-1, Tsu-0 and At2g21045 mutation were used in the study and a boxplot analysis was performed (Figure 4.20)

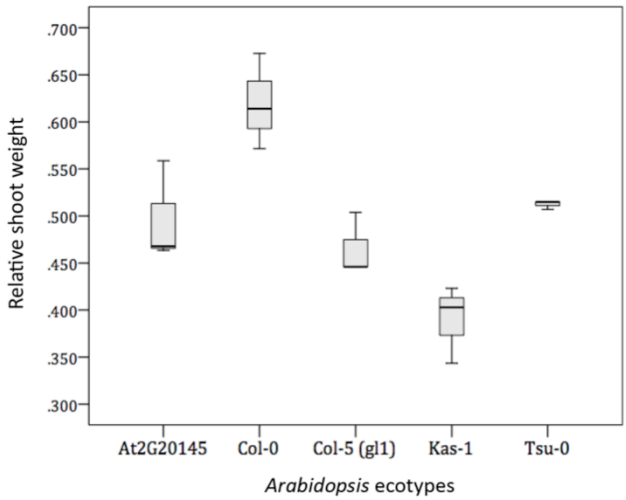

Figure 4.20. Comparing Col-0, Col-5(gl1), Kas-1, Tsu-0 and At2g21045 mutation using post-germination assay under 4ppm arsenite stress. Data are means of four replications (±SD).

A significant difference was found between the five ecotypes in One-way ANOVA (*F*-value=12.512, *P*-value=0.01). Using the same data, Post Hoc tests divided the five ecotypes into two groups at a significant level of 0.05. Col-0 ecotype belonged to one group on its own and was the most arsenite-tolerant ecotype. The other 4 ecotypes belonged to another group and had statistically similar sensitivity level. Kas-1 was the most arsenite-sensitive one among them.

In the study of As(V) (Sánchez-Bermejo et al. 2014), %RGR was a physiological parameter for arsenate tolerance. In our study, the relative shoot weight from the post germination assay was considered to be the physiological parameter for arsenite tolerance in this research. Comparing measurements from %RGR and the post germination assay in 25 *A. thaliana* genotypes, no correlation between two sets of

data could be found (Figure 4.21).

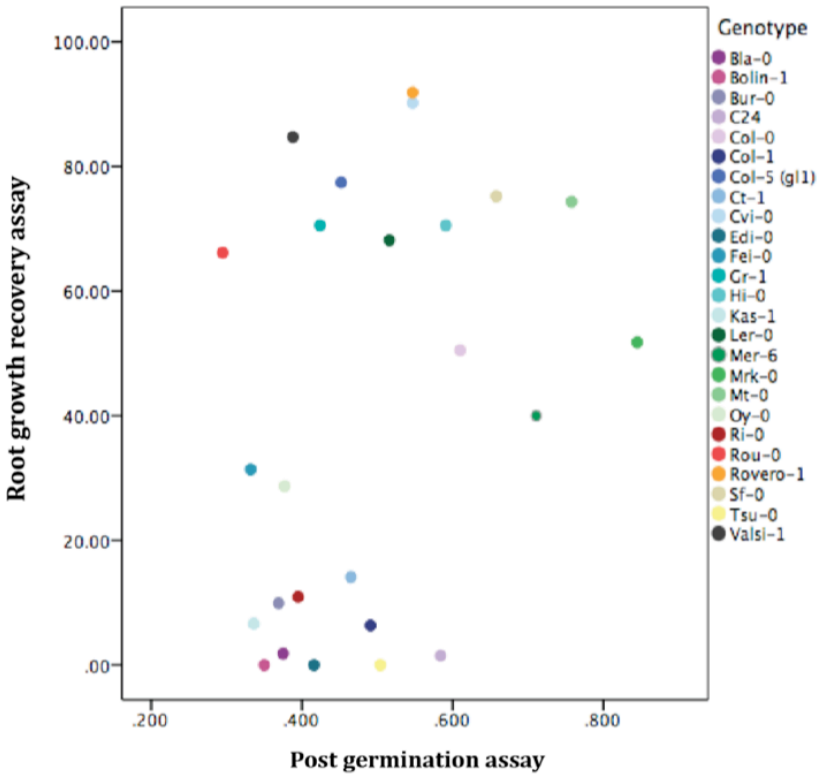

Figure 4.21. Correlation of the RGR assay and post-germination assay of 25 *A. thaliana* accessions.

4.4 Discussion

4.4.1 Role of anthocyanin in arsenite tolerance

A previous study has suggested that anthocyanins play a role in defense against As stress (Juszczuk et al. 2004). Under abiotic stress, anthocyanins are often observed in plants and normally there is a correlation between the accumulation of anthocyanins and the increasing stress level. Normally, anthocyanins are induced in plant vegetative tissues in response to abiotic stresses such as drought, high salinity, excess light, and cold (Chalker-Scott 1999), where they often correlate with enhanced stress tolerance (Zhu et al. 2013; Nakabayashi et al. 2014). A recent study found that different anthocyanins may play different roles depending on the kind of abiotic stress (Kovinich et al. 2015). During abiotic stresses, numerous roles of anthocyanins have been proposed, including functioning as ROS (Gould et al. 2002), photoprotectants (Tattini et al. 2014), and stress signals (Pourcel et al. 2013).

In this study, an accumulation of anthocyanins in seedling leaves under As stress was observed. Using eight *A. thaliana* genotypes, the concentration of anthocyanins is strongly correlated with the absolute shoot weight under 6ppm arsenite stress, as well as with the relative shoot weight change under the same arsenite stress. These results are consistent with previous studies, which is well documented that the *tt4* alleles contain mutations in the gene encoding CHS and, as there is only one single gene encoding this enzyme in *Arabidopsis*, *tt4* mutation is deficient in anthocyanins production (Saslowsky et al. 2000). If anthocyanin is involved in arsenite detoxification in any aspect, the *tt4* mutation should have different shoot weight compared with its background line Col-0. This was not observed in this research.

Even if anthocyanins have no detoxification contribution, they can provide a visible signal to indicate if plants are under arsenite stress when growing in an As polluted environment. Apart from anthocyanins accumulation, and shoot weight change, loss

of chlorophyll were found when a plant was under arsenite stress. However, shoot weight change and chlorosis can be caused by many reasons in field condition, therefore cannot be used as indicator for As accumulation. Anthocyanins can be tested using a variety of visualisation techniques over a wide range of growth stages (Peer et al. 2001) and can be used as a qualitative measure to indicate when a plant is under arsenite stress and accumulating As.

4.4.2 Arsenite reduced and delayed seed germination

To test the influence of seed size when seed is germinating in arsenite, the seeds weight of four genotypes were physically increased by decreasing the total amount of seed plants produced. Though this method showed more effect in genotype Mt-0 than another three genotypes, it significantly increased the seed weight and seed size. Perhaps this was due to genotype Kas-2 and genotype Bur-0 produce big seeds naturally, therefore reducing the amount of siliques would provide limited improvement. The relative change in both seed weight and size of genotype Bur-0 and genotype Kas-2 were only increased about 20%, respectively. The relative change of seed weight and size of genotype Ler-0 was 25% and 18%, respectively. The result could be due to the plant size of genotype Ler-0 is naturally short and small. The relative siliques number change was not very suggestive comparing with genotype Mt-0. The genotype Mt-0 had the largest change. This genotype is naturally tall and high yielding, thus reducing siliques number should significantly help the remaining siliques to obtain more nutrients and produce larger and heavier seeds.

Our results clearly indicate that larger seeds germinate better and faster under As stress. Increase in the biomass of seeds could increase both the germination speed and germination rate. After germination, other mechanisms help plants to defend themselves against arsenite stress. The reciprocal influence from both germination and the post-germination stage would cause error of measurement since arsenite has different effects on the two stages. This may be the reason for the big data deviation

that was observed in the measurement of root length change under arsenite stress (Figure 3.12)

Combining all results, it was found that if seeds are germinated and stay in a 4ppm arsenite environment, some seeds could not germinate or die after a certain period of time after germination. However, a plant can grow if it is transferred into a 4ppm arsenite environment after a 7-day seedling nursery. Indeed, the hydroponically cultivated plants grew under 4ppm arsenite, and could grow to bolting, flowering stage, and produce seeds, because they germinated and had seedling nursery stage without arsenite stress for 7 days. This observation verified the value of seedling nursery. For example, if a crop plant is sensitive to arsenite but provides the edible tissue with low As content, seedling nursery could increase its growth rate, therefore to produce an economic value along with the clean agronomic products.

4.4.3 Phenotypic change under arsenite stress

It was confirmed that larger seeds grow better under arsenite stress, but the plants start to lose the growth advantage when they start to develop true leaves. Based on all these observations, a post-germination assay was designed to measure the seedling arsenite tolerance ability. All seeds were grown in an environment for optimal germination.

Ler-0 was found to be the most arsenite tolerant type, whereas Kas-2 was the most sensitive type (Figure 4.16). Data collected from the 4ppm arsenite treatment and 6ppm arsenite treatment had a consistent result in the post-germination assay. Hence, Ler-0 and Kas-2 were selected as the mapping parents for genetic mapping. The interactions between genes from two genotypes and the environment together would cause a diverse plant phenotype. A mapping population using Ler-0 and Kas-2 as the parents with 154 lines available had been generated and published for mapping flower-time related QTLs (El-lithy et al. 2006). This population was chosen to perform

genetic mapping for arsenite tolerance in this research.

The shoot weight was selected as the trait for arsenite tolerance mapping. From the observations, arsenite stresses from both 4ppm and 6ppm could give a clear distinction of phenotypic change. Considering the difference of the genome might be less diverse in a RILs population than nature variations, the 6ppm arsenite as it was a higher stress was chosen for genetic mapping analysis.

The rosette diameter change under arsenite stress was evident, but the rosette diameter change does not correlate with the change of other phenotypic characters. Also no significant difference was found in relative rosette change under arsenite stress between various genotypes. Thus, the change of rosette diameter was not considered as an arsenite related parameter.

4.4.4 Analysis of AtPCS1 and ATQ1 under arsenite stress

Two genes have been reported involving in plants As biochemistry pathway by the time of this thesis been made. One of them was AtPCS1, which encoding a phytochelatin synthase to help plants to convert As^{III} to non-toxic form (Vatamaniuk et al. 1999). Previous studies have examined the function of the AtPCS1 gene in the resistance of heavy metal stress (Wojas et al. 2010). Another gene was discovered more recently, a study discovered an ATQ1 gene encoding an arsenate reductase in *A. thaliana* through gene mapping (Sánchez-Bermejo et al. 2014). Lacking the expression of this gene could caused a highly arsenate sensitive. In our study, it was necessary to eliminate the possibility that the phenotypic change perceived in shoot weight in the post-germination assay was due to the effect from these two genes.

Using the housekeeping gene as the internal control, RNA samples were extracted from post-germination assay treated samples. Genotypes Bur-0, Kas-2, Ler-0 and Mt-0 were analyzed because they had very different phenotypes. First, the expression

difference of the gene AtPCs1 could not be found between the arsenite treated samples and the control group. Second, the expression differences of the gene AtPCs1 between four tested genotypes approved. Therefore, it can be confirmed that the expression of this gene dose not changed under arsenite stress in posy-germination assay.

ATQ1 encodes an arsenate reductase, which has been well studied (Sánchez-Bermejo et al. 2014). As(V) needs to be reduced to As(III) and combined with chelate, and then sequestered in cell vacuoler. Using genetic mapping to study arsenate tolerance, the contribution of ATQ1 has been confirmed (Sánchez-Bermejo et al. 2014). However, in our study, arsenite was used, and in theory, this gene should not make any impact.

Figure 4.2 shows results of the study of arsenate (Sánchez-Bermejo et al. 2014). In the study, the labeled Col-0 is the Col gl-1, which is termed Col-5 on NACS website. Here in my study, the same T-DNA line, Col-5 and Kas-1 were analyzed in the post-germination assay to compare with relative growth rate change (%RGR) assay. The arsenite concentration used in this research was 4ppm arsenite which is almost equal to 30μM As(III). Combining the results from the published %RGR assay and post-germination assay, those genotypes that had a high tolerance to As(V) also had a high tolerance to As(III). Kas-1 was an exception showing more sensitivity to As(III) in post-germination assay than in relative growth rate change (%RGR) assay.

Using 25 genotypes, the correlation between %RGR assay and post-germination assay was compared. If the same mechanism is operative underlying the two measurements, the two data sets should correlate with each other. However, this correlation was not discovered (Figure 4.21). It is certain that plants have different tolerance mechanisms and show different phenotypes to cope with As(III) and As(V).

4.4.5 Summary

In this chapter, a new assay was developed to obtain the measurement of arsenite tolerance phenotype. A post-germination assay was designed based on the findings that larger seeds germinated significantly faster and better than smaller seeds under arsenite stress and seed size did not relate to seedling growth against arsenite stress, since the seedling emergence stage to compare multi-genotypes at post-germination stage.

The data of the shoot biomass weight change and the chlorophyll fluorescence change of seedlings between day-7 and day-14 was collected and analyzed. Merging results from chapter 3 and this chapter, Ler-0 was selected as the most arsenite tolerant type and Kas-2 was selected as the most arsenite sensitive type among 21 *A. thaliana* genotypes. A RIL mapping population from a cross between Ler-0 and Kas-2 was chosen for genetic mapping analysis for arsenite tolerance trait.

The anthocyanin fluorescence was tested for its function under arsenite stress in plants. When *A. thaliana* plants grow under arsenite stress, they produce anthocyanins in the leaves, but anthocyanins are not used for arsenite detoxification and probably are produced as by-products or stress signals from certain biochemical pathways.

Many previous researches have found genes such as AtPCS1 and ATQ1 can help plants to tolerate heavy metal stress, but these genes are not responsible for the variation of arsenite tolerance in this research. The variation of phenotypic characters change under arsenite stress among *A. thaliana* accessions need to be further studied.

Chapter 5

Genetic Mapping for Arsenite Tolerance Loci

5.1 Introduction

After examining phenotypic changes under arsenite stress, Ler-0 was classified as the tolerant genotype and Kas-2 was classified as the susceptible genotype. In this chapter, a RILs population from a cross between Ler-0 and Kas-2 was selected to carry out the genetic mapping.

The advantage of using a published mapping population is to use its obtainable genetic map. After generating a mapping population, the next step is to construct genetic map. To produce a genetic map, researchers collect DNA samples of all the lines in the population, in which a certain trait is prevalent, and then using designed molecular marker to examine it for unique genetic patterns that are seen in the mapping population. These characteristic patterns in the chemical bases that make up DNA are referred to as markers. This step together with the generation of mapping population can take up to 5 years depending on the species, and it may take 3 years in the research of *A. thaliana* RILs population (Silady et al. 2011). Generally, a published *A. thaliana* RIL mapping population would come with a genetic map, which was developed by at least 80 markers. Thus, using a published mapping population would not just be time saving, but also cost effective. The only disadvantage of using a published mapping population is its fixed mapping resolution, due to the fixed number of markers.

The Ler-0 x Kas-2 RILs is one of three populations belonging to a multi-parent mapping population (El-lithy et al. 2006). In this population, Ler-0 was selected as a common mother to cross with Antwerp-1, Kashmir-2, and Kondara (El-lithy et al. 2006). Figure 5.1 is the genetic map of Ler-0 x Kas-2 RILs. Most of the genetic markers used to construct this map were first selected from SNPs between the reference accessions Col-0 and Ler-0 and then used to genotype all lines of Ler-0 x Kas-2 RILs. The distance between adjacent markers was less than 13 cM (El-lithy et al. 2006).

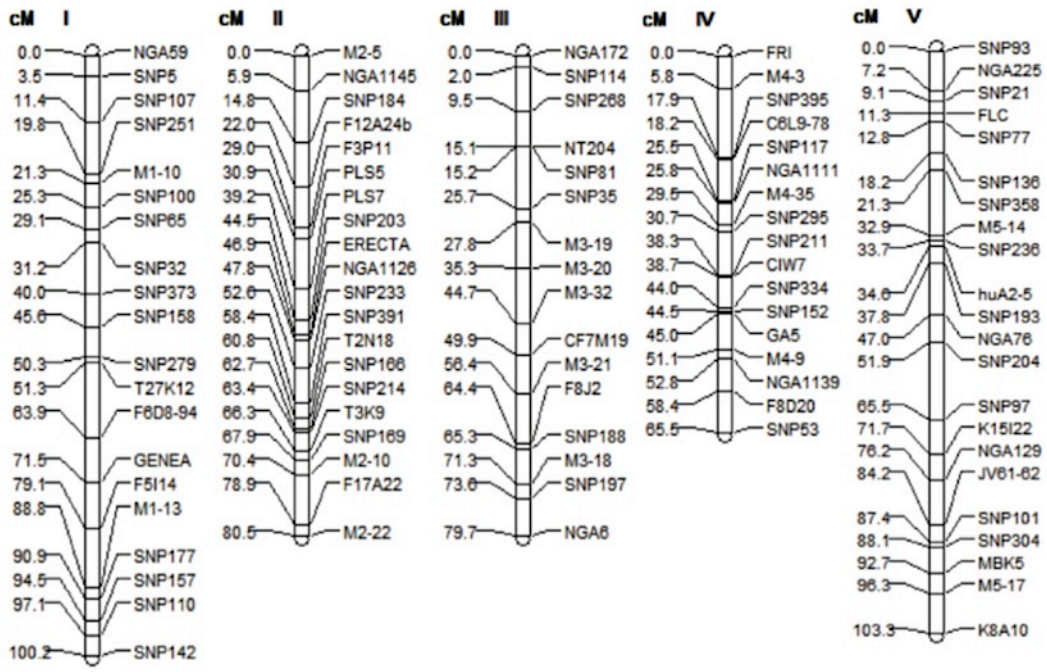

Figure 5.1. The genetic map of Ler-0 x Kas-2 RILs mapping population (El-lithy et al. 2006).

Association mapping was used as a complementary alternative to the bi-parental mapping to improve mapping resolution in this research. When using published genotype data of a bi-parental mapping population, the genotype information restricts the mapping resolution because of the fixed marker linkage distance. Association mapping can also incorporate the effects of the classical bi-parental population mapping to reveal the genetic basis of the complex traits (Lou et al. 2006; Segura et al. 2012). The association mapping is based on the extent of Linkage Disequilibrium (LD) observed in a set of accessions that are not closely related (Jorde 2000). In contrast to classical bi-parental population mapping, the association mapping detects relationships between phenotypic variations and gene polymorphisms in the existing germplasm. The increasing availability of genomic polymorphism data allows the more frequent use of genome-wide association studies (GWAS) to investigate the genetics of quantitative traits (Wang et al 2005). Normally, GWAS are carried out using a single-locus test, which simultaneously queries the relationship between the phenotype and the single-locus allele and their linear combinations in either

case-control populations or cohorts (Roeder et al. 2005).

Fine mapping, by adding more markers to a fixed genetic map, is another approach to increase the mapping accuracy (Rant et al. 2013). Genotyping methods, including Restriction Fragment Length Polymorphisms (RFLPs), Amplified Fragment Length Polymorphisms (AFLPs), Microsatellites or Short Tandem Repeat Sequences (STRs), and Single Nucleotide Polymorphisms (SNPs) can be applied in both whole genome genetic mapping and fine mapping analysis (Powell et al. 1996; Rafalski 2002). Because the region in fine mapping study is generally much smaller than in the whole genome QTL mapping, SNPs are generally more accepted marker types owing to their abundance (Storm and Darnhofer-2003). SNPs can be defined as bi-allelic variants within a population occurring with an allelic frequency higher than 1% (Jurinke et al. 2004). In SNP analysis between two accessions, the primer-binding site is placed adjacent to the polymorphic position. A specific extension product is generated for each allele, thus both elongation products can be distinguished in mass by one nucleotide. The two SNP alleles appear as two distinct mass signals and can be measured in a Matrix-assisted laser desorption/ionization time-of-flight mass spectrometer (MALDI-TOF Mass Spectrometry), which can either be a high-throughput or middle-throughput way to examine biomolecules of the amplified SNP marker fragment (Jurinke et al. 2004; Blievernicht et al. 2007). This assay can be performed directly from genomic DNA to genotype a case-control population or the natural accession (Jurinke et al. 2004). In this research, the fine mapping was carried out using MALDI-TOF mass spectrometry-based SNP genotyping.

5.2 Materials and Methods

5.2.1 Materials

All *A. thaliana* genotypes used in this chapter were listed in Section 2.1.1.2.

5.2.2 Main methods

The major method used in this chapter were:

Post-germination assay of *A. thaliana*;

Chlorophyll content of *A. thaliana* seeding;

DNA extraction;

QTL mapping of Kas-2 x Ler-0 RI population;

Fine mapping analysis;

GWAS analysis;

Searching database for arsenite tolerance related genes.

The detailed procedures of these experiments were given in Chapter 2.

5.3 Results

5.3.1. Measuring arsenite tolerance of Ler-0, Kas-2 and their RILs

The variation of shoot weight of all lines in the mapping population in response to arsenite stress was measured. The absolute shoot weight and the weight change under arsenite stress were used as the phenotypic measurements for the genetic mapping. The shoot weights of two parental genotypes Ler-0 and Kas-2, as well as their progeny population (151 RILs) were evaluated by post-germination assay simultaneously. Figure 5.2 illustrates the variance, which was probably caused by the QTL variation of arsenite tolerance in the presence of 4ppm arsenite. Figure 5.2.c shows the growth of the arsenite-tolerance genotypes, while Figure 5.2.d shows the growth of the arsenite sensitive genotype.

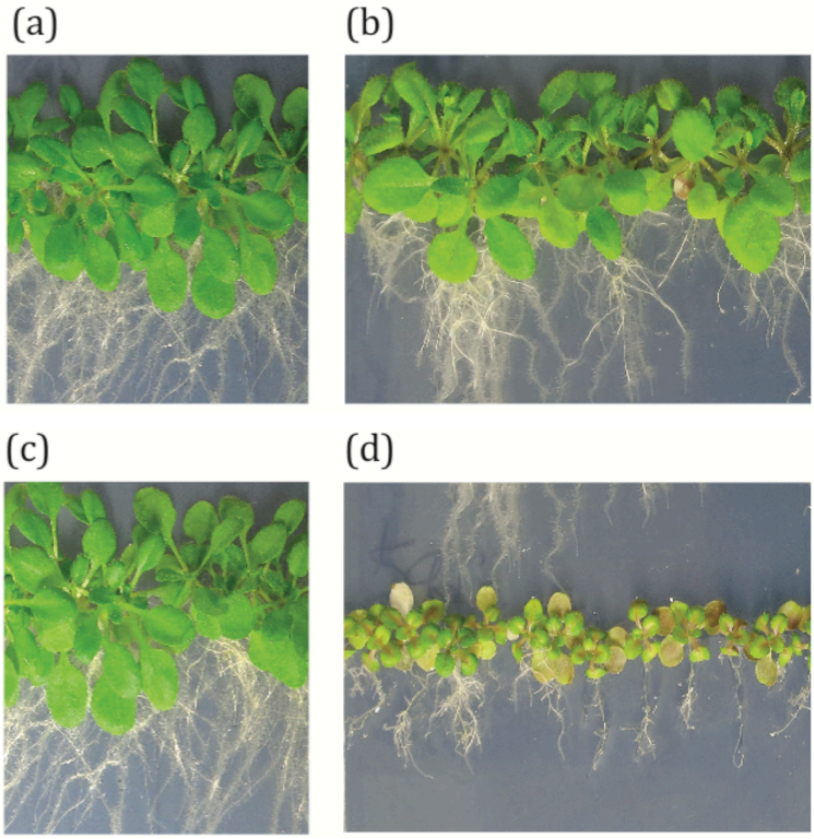

Figure 5.2. Representative images of 14-day old seedlings in control condition and under a 4ppm arsenite stress. a. Ler-0, 0ppm; b. Kas-2, 0ppm; c. Ler-0, 4ppm, and d. Kas-2, 4ppm.

Broad sense heritability (H^2) in this study is defined as the proportion of variance explained by between-line differences using the general linear model module. The variance was calculated using the general linear model module to measure 6 plant seedlings of each genotype. Measurements were repeated four times.

The result of the Shapiro-Wilk test indicated that the distribution of the absolute shoot weight of RILs was almost parametric (*P*-value=0.000) and can be used in QTL analysis. Two-way ANOVA showed a significant genetic effect (*P*-value=0.000) and a significant environment effect (*P*-value=0.000) on absolute shoot weight. Heritability of absolute shoot weight in Ler-0 x Kas-2 RILs population was 97%. The frequency distribution of mean absolute shoot weight of RILs population is given in figure 5.3.

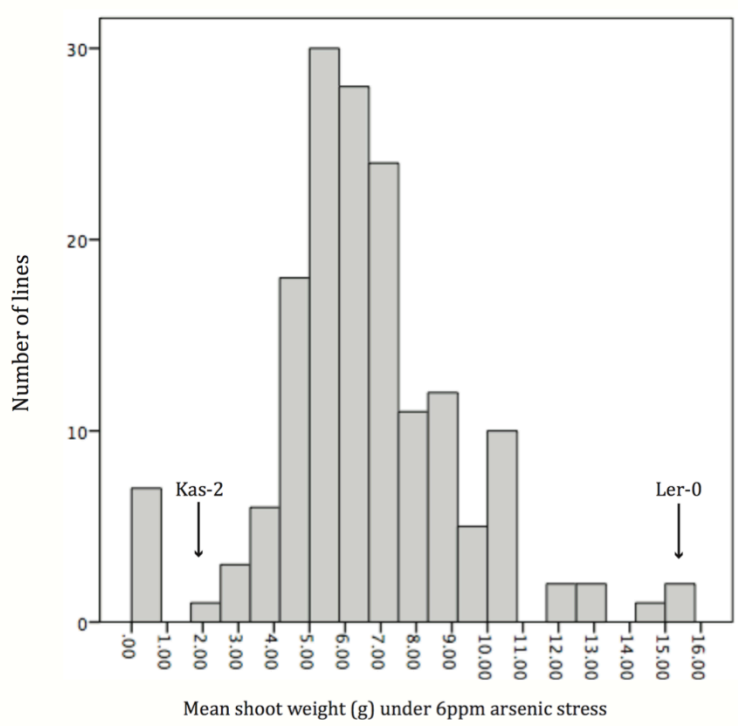

Figure 5.3. Frequency distribution of mean absolute shoot weight of Ler-0 x Kas-2 RILs in the presence of 6ppm arsenite. Arrows indicate the mean shoot weights for the parent lines Kas-2 and Ler-0.

The Shapiro-Wilk test indicated the data distribution of the relative change of the shoot weight was almost parametric (*P*-value=0.000). One-way ANOVA analysis showed the significant difference between RILs (*P*-value=0.000). Heritability of

relative shoot weight in Ler-0 x Kas-2 RILs population was 90.9%. The frequency distribution of mean relative arsenite tolerance of RILs population is given in figure 5.4.

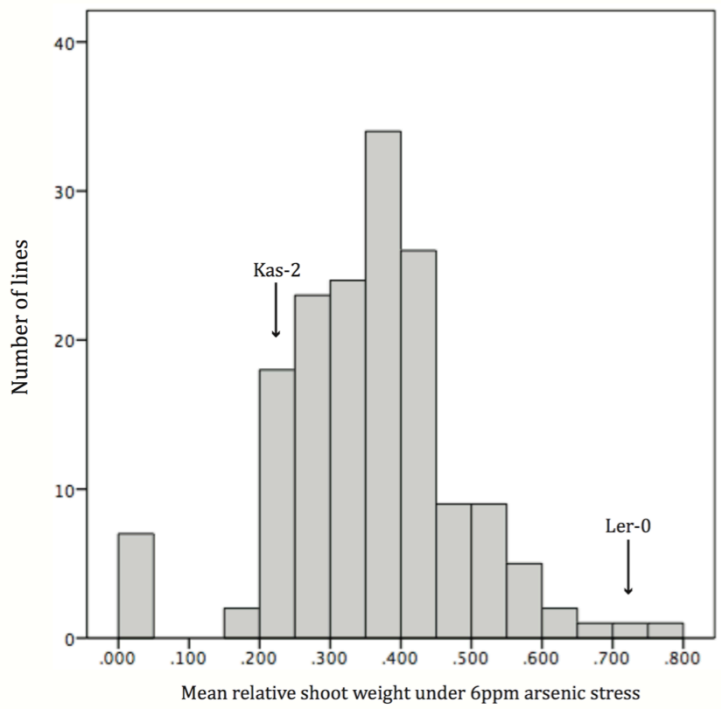

Figure 5.4. Frequency distribution of mean change of shoot weight by As^{III} of Ler-0 x Kas-2 recombinant inbred lines. The mean shoot weights for the parent lines are indicated by arrows.

5.3.2 QTL mapping using absolute shoot weight

The measurement of the absolute shoot weight of 151 RILs under 6ppm arsenite stress was used for both Interval Mapping (IM) (Figure 5.5) and Composite Interval Mapping (CIM) (Figure 5.6) analysis.

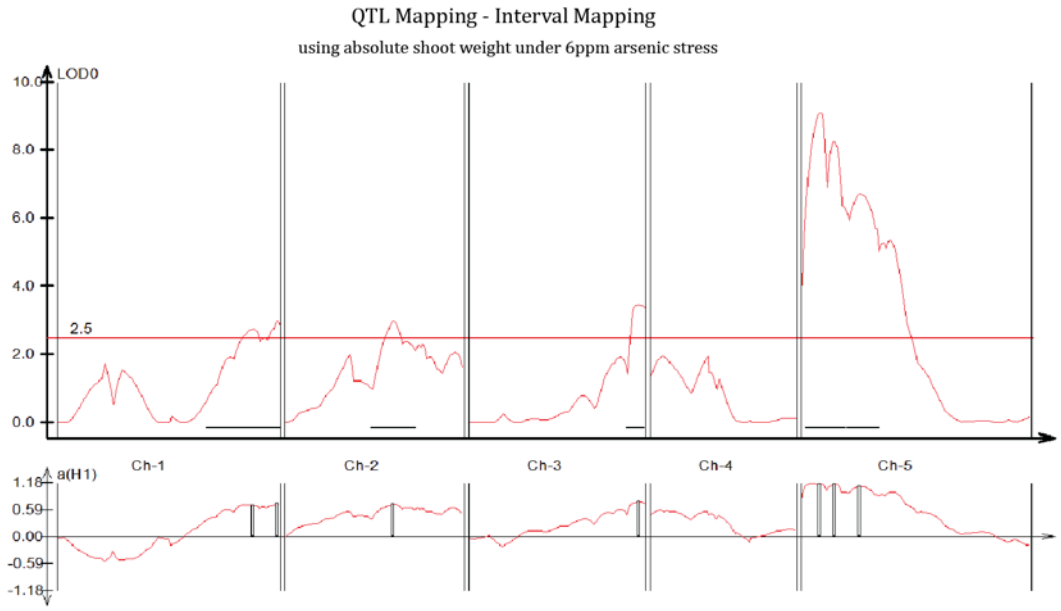

Figure 5.5. IM mapping result with the additive effect data of the absolute shoot weight under 6ppm arsenite stress.

Seven QTLs were detected in the IM analysis. The information of the QTLs location on the chromosome and on the linkage map (column **Position**), the amount of the effect contributed by each QTL (column **Additive effect**), and the percentage of each QTL (column $\mathbf{R^2}$) in the total genetic effect are summarized in Table 5.1. The mean absolute shoot weight of the 151 RILs is 6.87mg. The unit of additive effect is the same as the measurement of the trait.

Table 5.1 Detected QTLs from IM analysis using absolute measurement.

QTL	Chromosome	Position	Additive effect (mg)	R^2
1	1	0.8811	0.679	0.081688
2	1	0.9911	0.7207	0.092616
3	2	0.4881	0.7115	0.090858
4	3	0.7661	0.7623	0.103928
5	5	0.0821	1.1599	0.241245
6	5	0.1481	1.1578	0.240102
7	5	0.2631	1.1001	0.217239

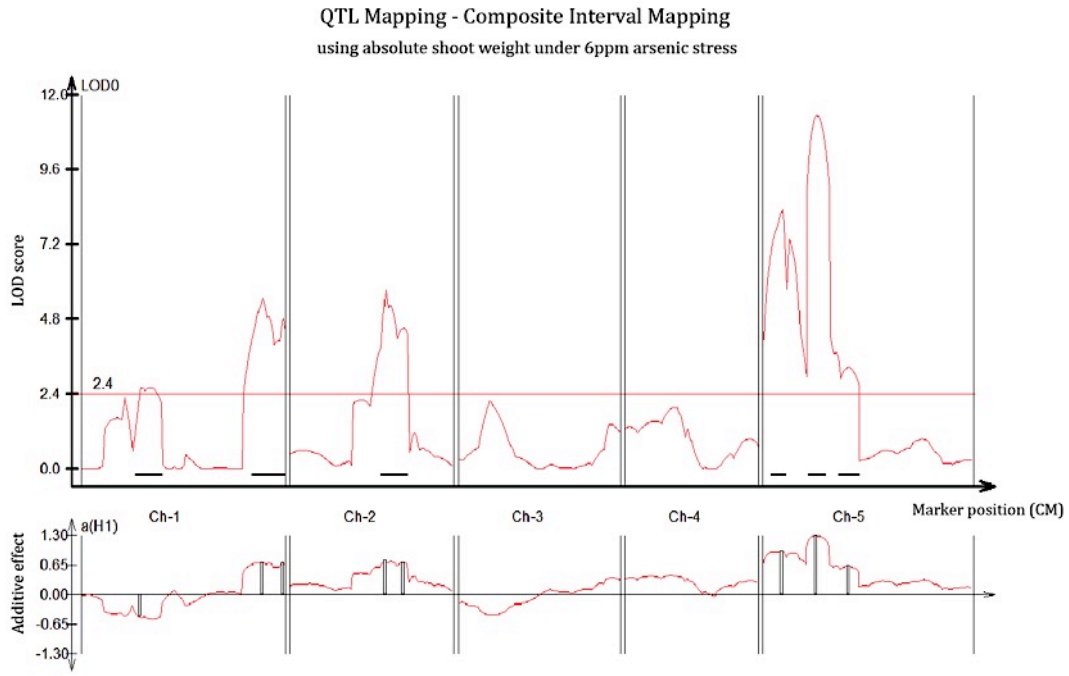

Figure 5.6. CIM mapping result with the additive effect data of the absolute shoot weight under 6ppm arsenite stress.

CIM analysis identified three loci on chromosome 5, which were also found in IM analysis. The only difference of the loci between IM and CIM analysis was the R^2 value. Information about QTLs is presented in Table 5.2.

Table 5.2 Detected QTLs from CIM analysis using relative shoot weight measurement.

QTL	Chromosome	Position	Additive effect (mg)	R^2
1	1	0.8811	0.679	0.081688
2	1	0.9911	0.7207	0.092616
3	2	0.4881	0.7115	0.090858
4	3	0.7661	0.7623	0.103928
5	5	0.0821	1.1599	0.241245
6	5	0.1481	1.1578	0.240102
7	5	0.2631	1.1001	0.217239

5.3.3 QTL mapping using relative shoot weight

The relative change of shoot weight of 151 RILs and their parent lines under 6ppm arsenite stress was measured by a post-germination assay, and it was then analyzed by IM and CIM analysis. Result of IM analysis is shown in Figure 5.7 and CIM analysis is shown in Figure 5.8.

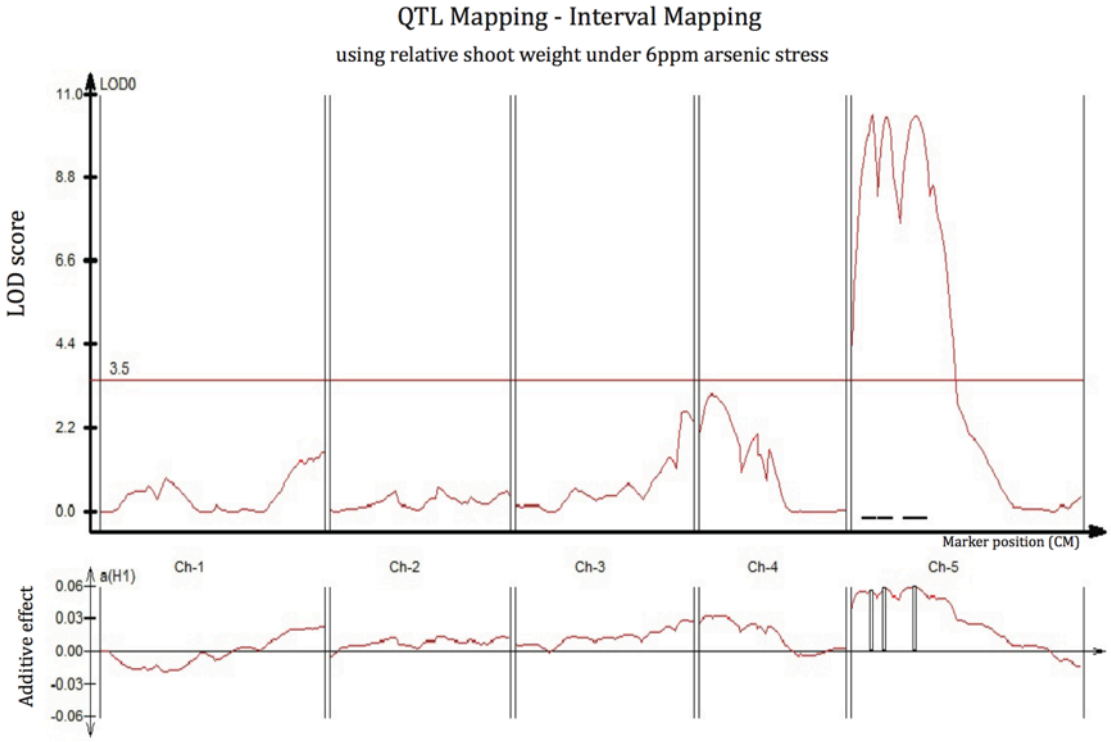

Figure 5.7. IM mapping result with the additive effect data of the shoot weight change data caused by 6ppm arsenite stress (relative shoot weight).

Three QTLs were detected in the IM analysis from the relative data of shoot weight. All three QTLs were located on chromosome 5, and they were near the location found from the IM analysis and the CIM analysis of absolute shoot weight data. The information, including OTLs location on chromosome and its position on linkage map, the amount of the effect contributed by each QTL, and the percentage of each QTL of the total genetic effect is given in Table 5.3. Mean trait measurement of relative shoot weight change under 6ppm arsenite stress was 0.37, which was the mean fraction from the arsenite treated shoot weight against control group shoot weight of every line of all 151 RILs. The unit of additive effect from the loci is the same as the trait quantity, which is also a fraction.

Table 5.3 Detected QTLs from IM analysis

QTL	Chromosome	Position	Additive effect	R^2
1	5	0.0911	0.056	0.270767
2	5	0.1481	0.0582	0.292548
3	5	0.2831	0.0605	0.316795

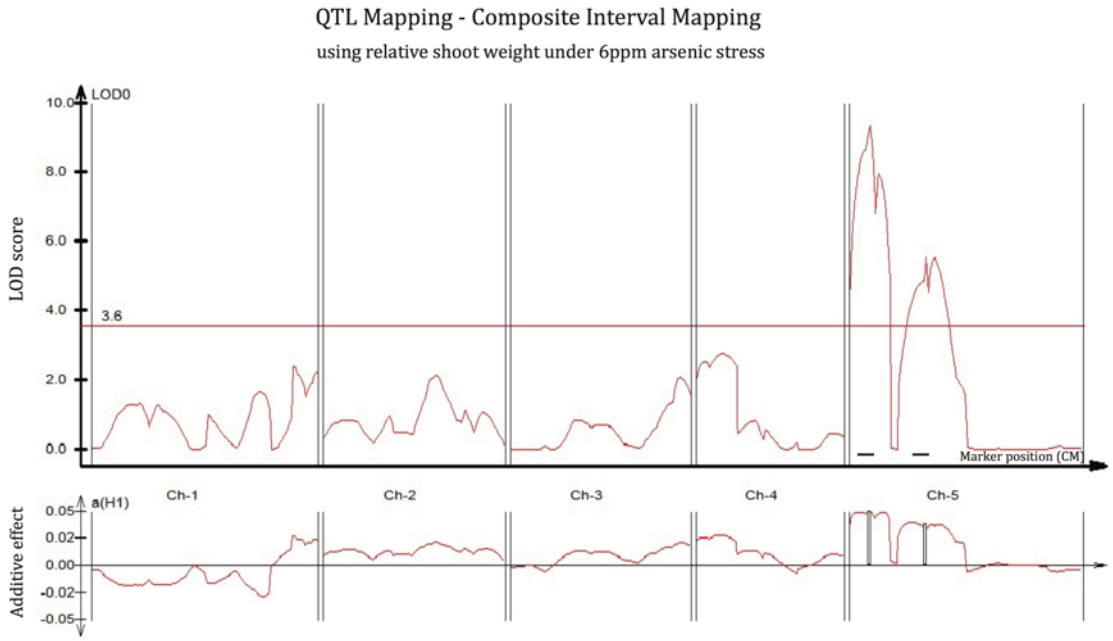

Figure 5.8. CIM mapping result with the additive effect data of the shoot weight change data caused by 6ppm arsenite stress (relative shoot weight).

One locus found from the IM analysis was also detected in the CIM analysis. However, the information of rest loci found in the IM analysis was not consistent with in the CIM analysis. Information of the loci found from the CIM analysis is presented in Table 5.4

Table 5.4 Detected QTLs from CIM analysis

QTL	Chromosome	Position	Additive effect	R^2
1	5	0.0911	0.047	0.174092
2	5	0.3371	0.0363	0.09825

The Two-LOD analysis was used to examine the locus information to find the area that is more likely to contain QTLs on the chromosome 5. The Two-LOD supporting intervals were established at 95% QTL confidence interval level. A region between 1.3mb-7.6mb on *A. thaliana* chromosome 5 was located, which should contain the major effect loci underlying arsenite tolerance.

5.3.4 Design of the fine mapping markers and the genetic map

A fine mapping analysis followed by adding more markers in the existing genetic map of Ler-0 x Kas-2 population to increase the mapping resolution was carried out in order to narrow down the intervals of the QTLs. The insertion-deletions (InDels) marker and SNP markers, which are both dependent on the variation between the genetic sequences of Ler-0 and Kas-2, were both considered and used for fine mapping genotyping.

InDels markers were selected because of the availability of InDels information of Ler-0 and Kas-2, which made it possible to use the different sequence length-based genotype. Also, using electrophoresis can analyze the different sequence length based genotype easily. Six pairs of InDels markers were designed, which in theory, can provide the different length of fragments of Ler-0 and Kas-2 from PCR products. The result is shown in Figure 5.9, but none of the markers showed a clear difference between Ler-0 and Kas-2.

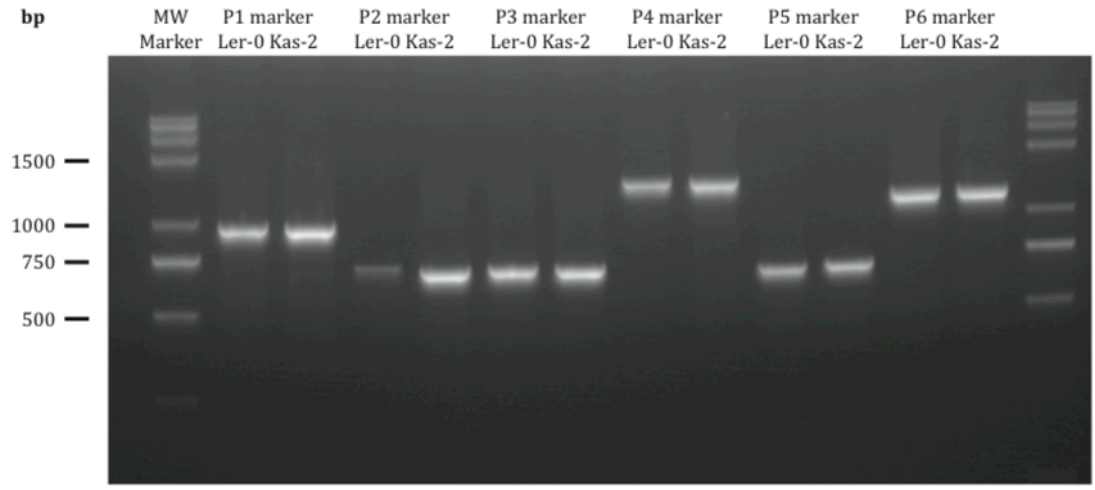

Figure 5.9. Electrophoresis result of InDel markers selection for genotyping Ler-0 x Kas-2 mapping population is presented. Results from the use of 2% MS agarose gel was used.

Since InDels marker could not explore the genotype variation between accession Ler-0 and Kas-2, SNP difference was used for marker selection. After comparing the SNP difference between Ler-0 and Kas-2 in the target region on chromosome 5, 30 SNPs on chromosome 5 were selected. For genotyping markers selection, the distance between each marker was designed at 0.1mb. Figure 5.10 shows the distributions of the 30 SNP markers together with another 8 markers from published research from El-lithy et al. (2006).

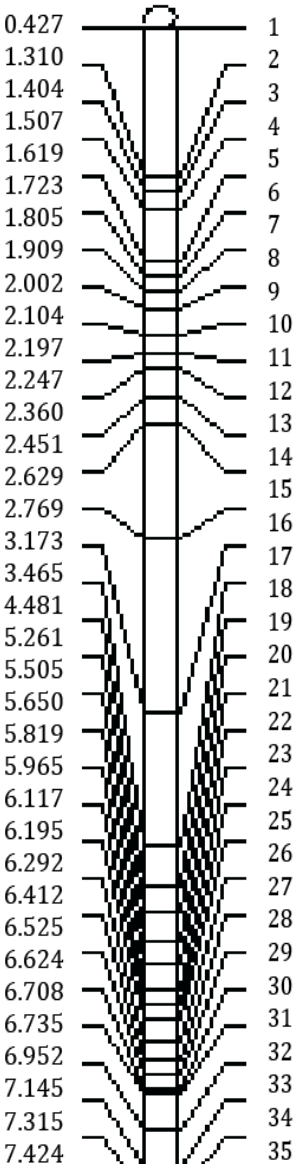

Figure 5.10. The physical distribution of SNP markers on Chromosome 5.

MALDI-TOF mass spectrometry-based SNP genotyping was used to analyze the allelic variation between accession of Ler-0 x Kas-2 at the SNP marker location. The genotype data of tested RILs are given in supplementary data 2. The genotyping data and the SNP marker physical location were processed by Joinmap 4.0 to draw a new genetic map (Figure 5.11). The genetic map contained 38 molecular markers.

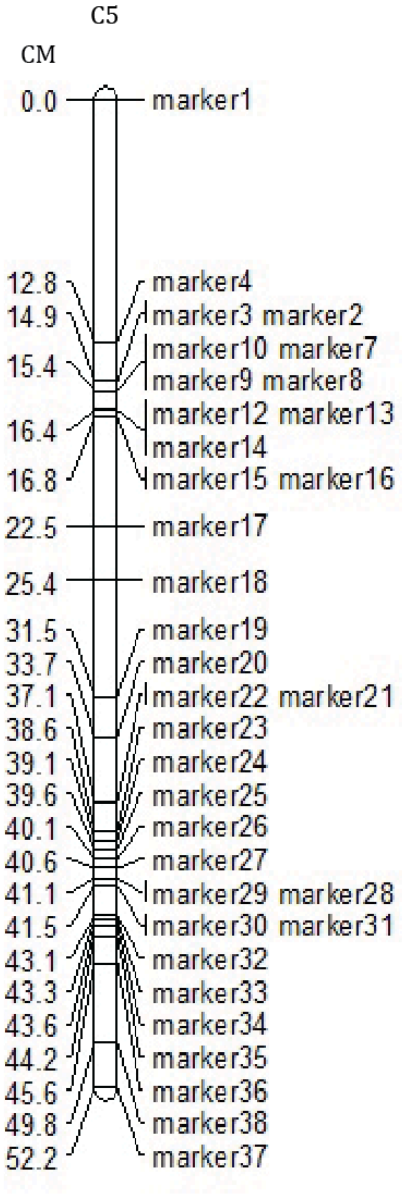

Figure 5.11.The linkage distance of 38 molecular markers was calculated and the linkage map was drawn using Joinmap 4.0

5.3.5 Fine mapping using relative shoot weight

Using relative shoot weight data, and the genetic map, fine mapping analysis was performed through the IM and the CIM analysis. The result of the IM analysis is given in Figure 5.12 and the result of the the CIM analysis is in Figure 5.13. Using relative shoot weight, the QTLs were found in both the IM and CIM analysis. The trait measurement of relative shoot weight was the fraction of the arsenite treated shoot

weight data against control group shoot weight and the mean value is 0.37. The QTLs information is given in Table 5.5 and Table 5.6.

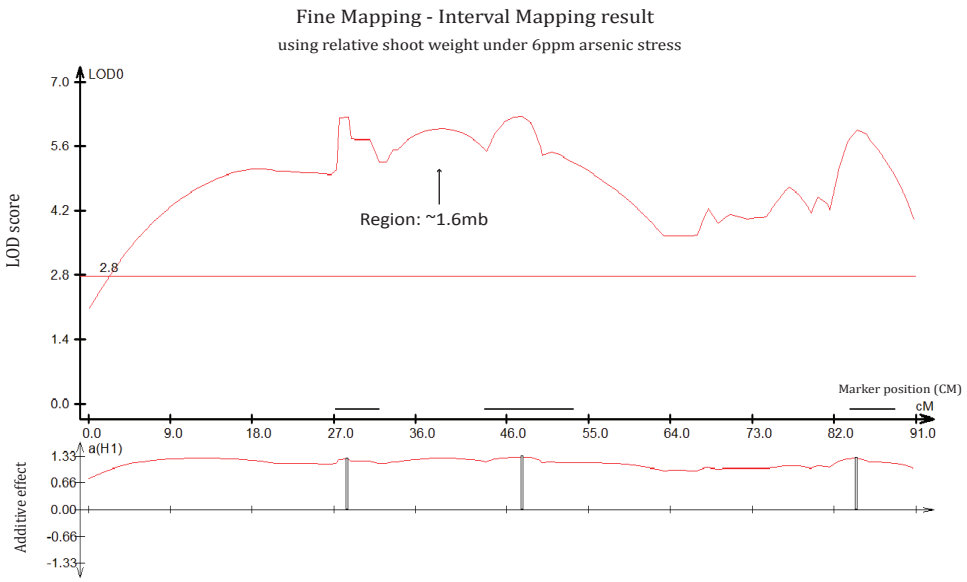

Figure 5.12. IM mapping result using the shoot weight change data caused by 6ppm arsenite stress (relative shoot weight).

Table 5.5 Detected QTLs from fine mapping IM analysis

QTL	Chromosome	Position	Additive effect	R^2
1	5	0.2851	0.0588	0.248646
2	5	0.4671	0.0636	0.29255
3	5	0.8451	0.0677	0.332932

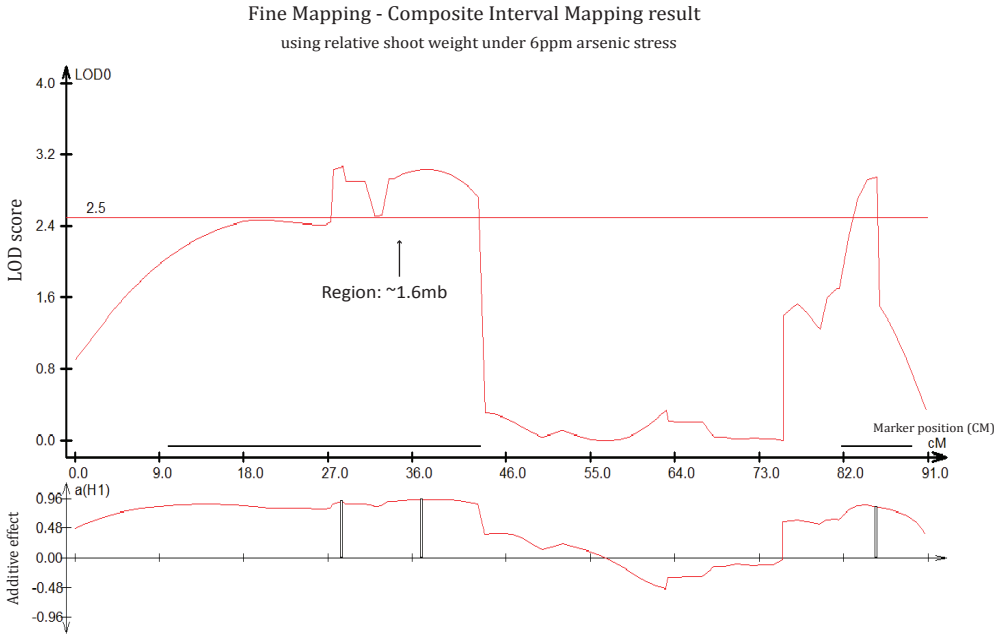

Figure 5.13. CIM mapping result using of the shoot weight change data caused by 6ppm arsenite stress (relative shoot weight).

Table 5.6 Detected QTLs from fine mapping CIM analysis

QTL	Chromosome	Position	Additive effect	R^2
1	5	0.3801	0.0417	0.108223
2	5	0.8451	0.0452	0.118245

The analysis of absolute shoot weight showed consistent result with relative shoot weight, therefore, is not provided here.

5.3.6 Single marker analysis using fine mapping physical map

When using single marker analysis, the linkage between markers is not considered in the analysis. The physical map allowed the estimate of the highest LOD score of QTL position to be directly linked to the marker position. The LOD score peak location is

intuitively shown in the relative data analysis (Figure 5.14).

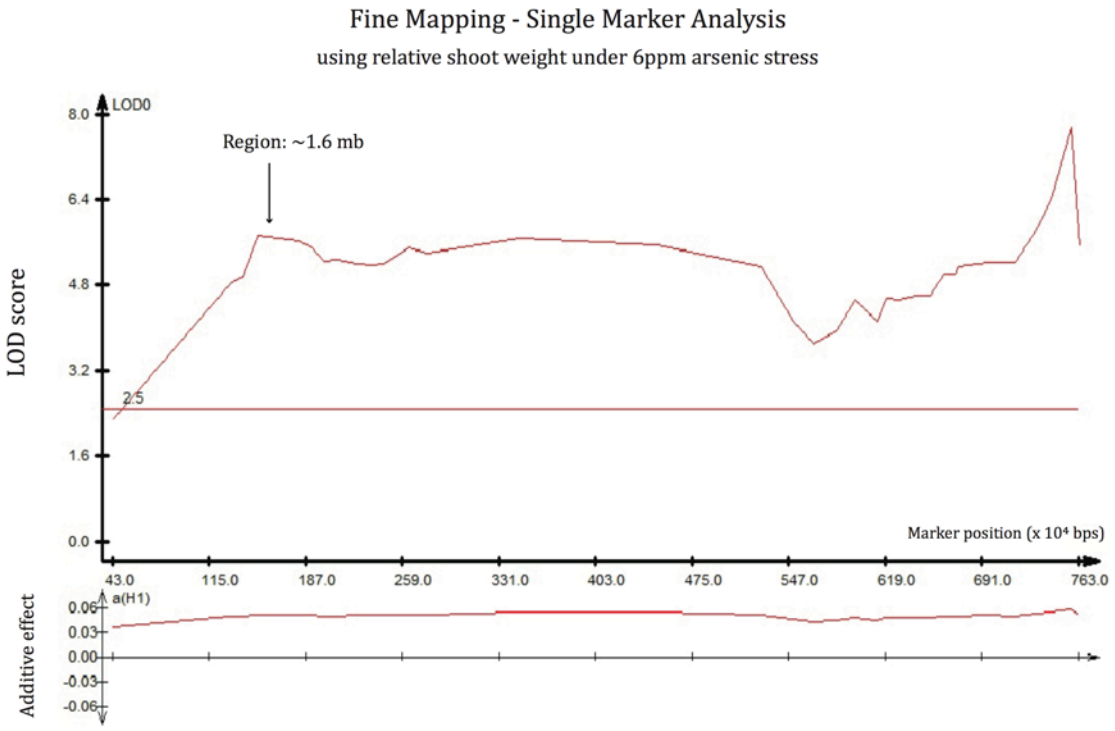

Figure 5.14. Single marker analysis result using the shoot weight change under 6ppm arsenite stress (relative shoot weight) with the additive effect data.

After integrating all the mapping results from IM analysis, CIM analysis and single marker analysis from the absolute data analysis and the relative data analysis in both general QTL mapping and a fine mapping, all the results were analyzed together and found one QTL located between the 1.5mb to 1.7mb region and another one located between the 4.75mb to 5.5mb on the chromosome 5 of the genome.

.

5.3.7 Arsenite tolerance of F1 hybrid plants

Because Ler-0 and Kas-2 mapping population is the RILs population, analyzing the genetic information of all the line of RILs in the QTLs analysis only obtained additive effect. A F1 heterozygote from inter-cross between Ler-0 and Kas-2 and its parents were examined to get more information about the effect. The boxplot figure given in Figure 5.15 shows the relative shoot weight change of Ler-0 and Kas-2 and their progeny under 4ppm arsenite stress.

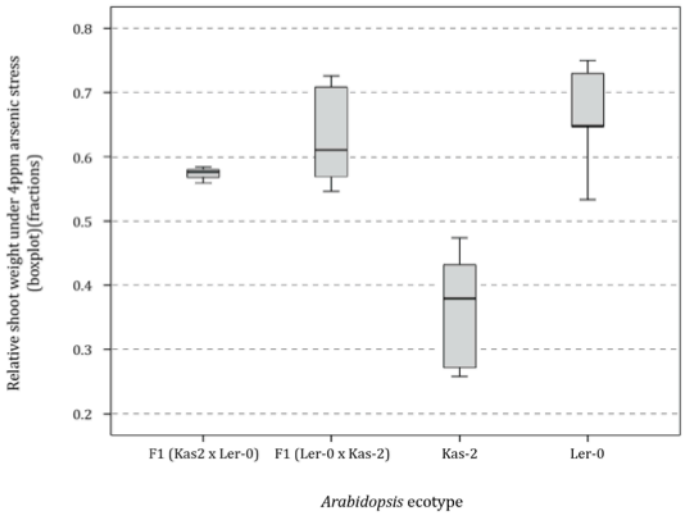

Figure 5.15. The boxplot of the shoot weight change caused by As^{III} (relative shoot weight) of Ler-0, Kas-2 and the F1. Data are calculated from four replications (±SD).

Homogeneous analysis classified relative shoot weight trait of Kas-2, Ler-0 and their F1 into two subsets. Kas-2 (mean=0.3655) was one group. F1 (Kas-2♀, mean=0.5731), F1 (Ler-0♀, mean=0.6320) and Ler-0 (mean=0.6926) had similar phenotype and they were another group. The size of the dominance effect is estimated as the difference from the parent mid-value. The figure seems to indicate that the Ler-0 allele is quite dominant over the Kas-2 allele.

5.3.8 GWAS analysis

Using the shoot weight change and chlorophyll content change of 39 genotypes under 4ppm sodium arsenite stress, a GWAS analysis was performed. The phenotypic data analysis was done by Dr. A. Brennan (Durham University).

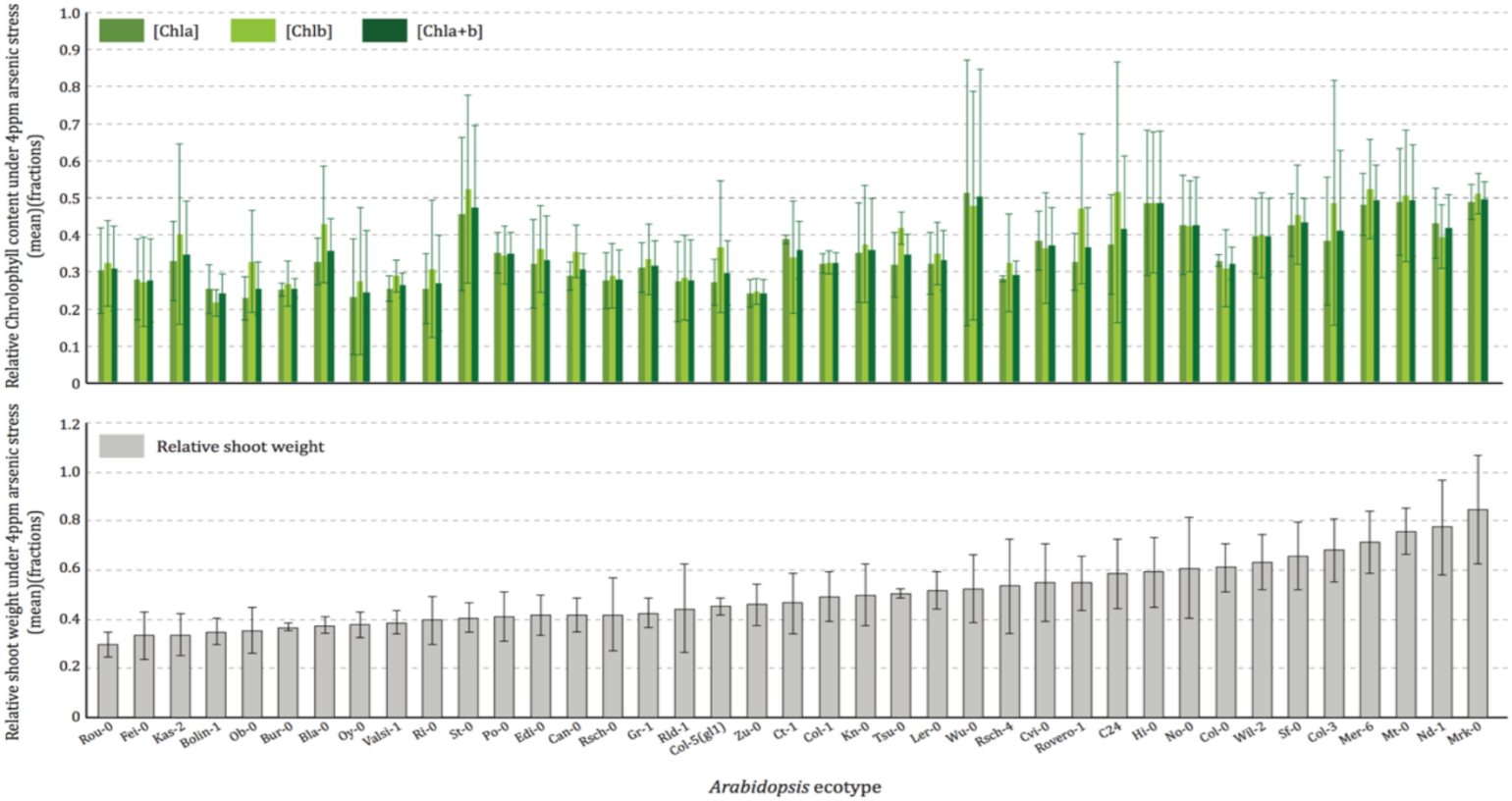

Figure 5.16. a. The relative chlorophyll content change caused by arsenite (the chlorophyll content of arsenite treated group: the chlorophyll content of the control group). Chlorophyll was extracted from 14 days old seedlings of 39 *A. thaliana* genotypes under 6ppm sodium arsenite stress. Data are means of four replicates (±SD); b. Relative shoot weight change caused by arsenite (the shoot weight of arsenite treated group: the shoot weight of the control group). Measurement was collected from 14 days old seedlings of 39 *A. thaliana* genotypes under 6ppm sodium arsenite stress. Data are means of four replicates (± SD).

Figure 5.16 contains the phenotypic data of all 39 genotypes. These data were arranged from the left to the right by the measurement of the relative shoot weight change. Using One-way ANOVA, a significant difference of relative shoot weight was found among genotypes (*F*-vaule=6.017, *P*-value=0.000). One-way ANOVA also confirmed a significant difference in the relative chlorophylls content change (Ct: *F*-value=8.409, *P*-value=0.000; Ca: *F*-value=6.277, *P*-value=0.000; Cb: *F*-value=9.539, *P*-value=0.000).

The GWAS analysis is presented in a Manhattan plot. Figure 5.17 is the result of the relative shoot weight change of 39 genotypes under 4ppm arsenite stress. Figure 5.18 is the result of the relative change of chlorophyll extracted from the shoots of 39 genotypes under 4ppm arsenite stress. The tolerance and chlorophyll traits did not suggest the same SNPs or regions. For the tolerance analysis, the best method seems to be GWAS with genetic principal components as cofactors. This method gave 4 out of the top 10 most significant SNP hits in the same region of chromosome 5 from 1623000 to 1658000 bp (personal communication with Dr. A. Brennan).

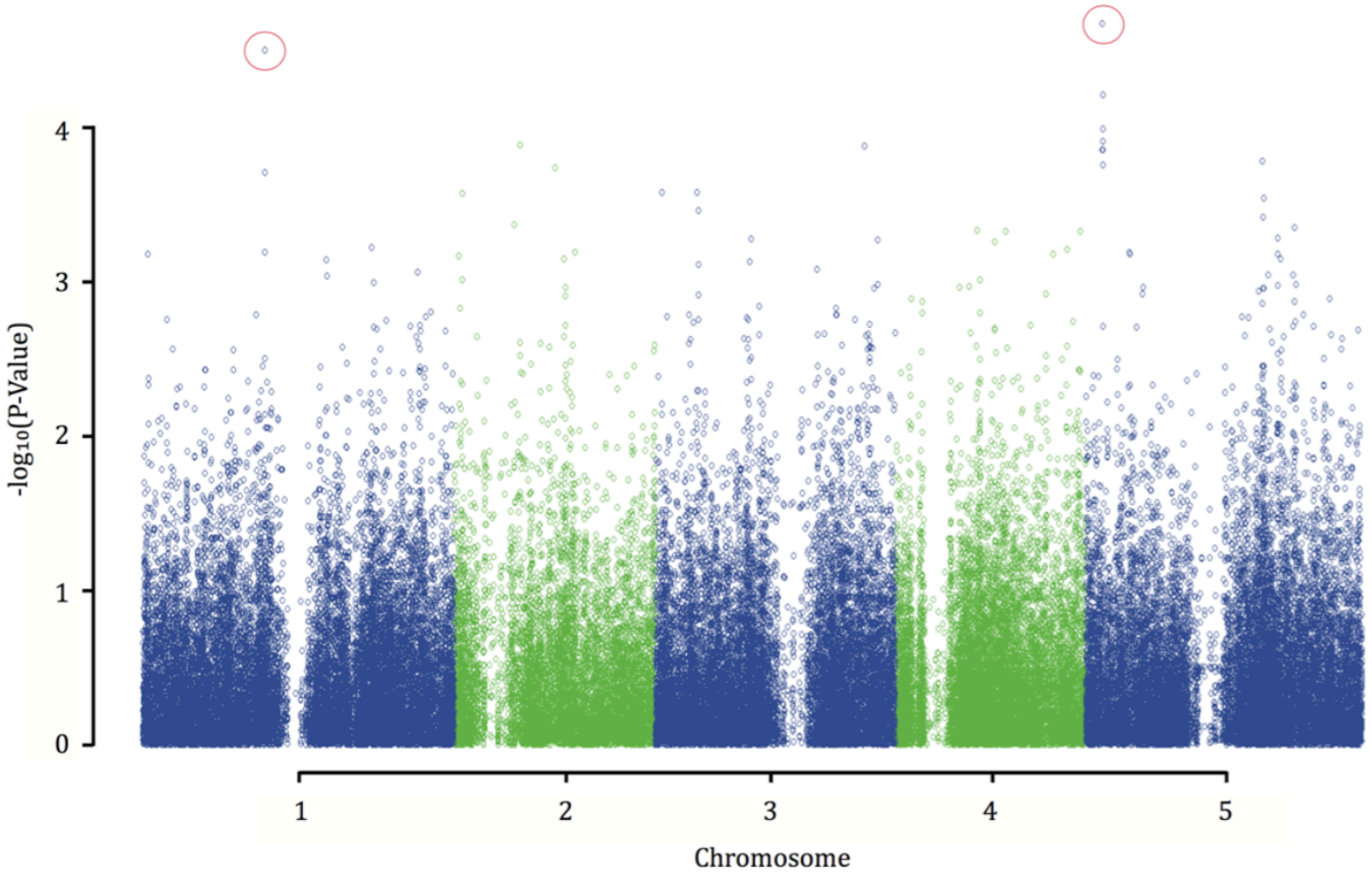

Figure 5.17. Manhattan plot of the relative shoot weight change of 39 genotypes under 4ppm arsenite stress

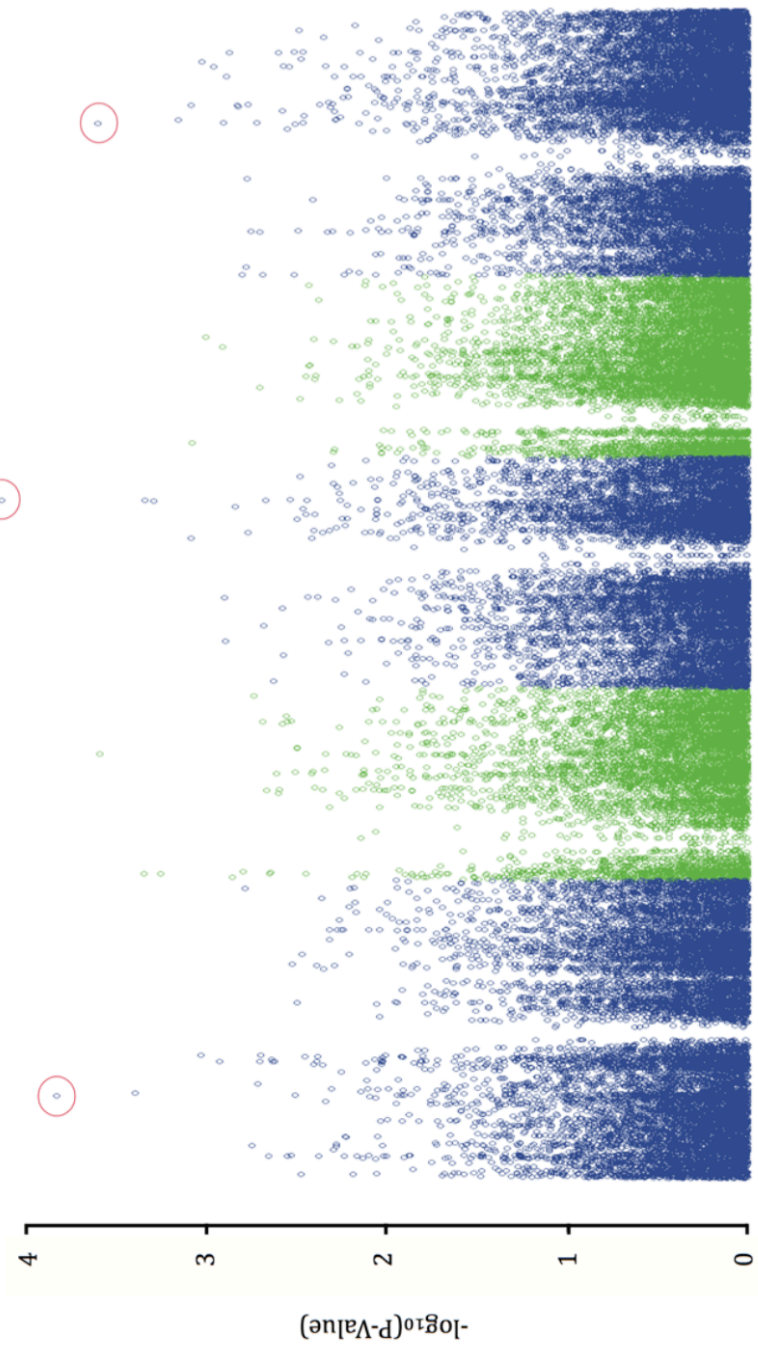

Figure 5.18. Manhattan plot of the change of relative chlorophyll extracted in shoot of 39 genotypes under 4ppm arsenite stress.

In chapter 4, an arsenite concentration of 6ppm was used to test the shoot weight change and the chlorophyll content change in 22 genotypes. In this chapter, the tested genotypes were extended to 39 genotypes, whereas the 22 genotypes used in chapter 4 were also included.

Comparing the shoot weight change and chlorophylls content under 4ppm and 6ppm arsenite stress, the shoot weight change under two different concentrations was highly correlated. A weak link between shoot weight and chlorophylls content under 4ppm arsenite stress was observed. There was no correlation between shoot weight change and chlorophylls content under 6ppm arsenite stress, and no correlation was found between chlorophylls content under either 4ppm or 6ppm arsenite stress. Information about the analysis is given in Table 5.7.

Table 5.7 Correlation between shoot weight and chlorophyll content under two different arsenite concentrations.

	Pearson correlation	*P*-value
Relative shoot weight under 4ppm arsenic stress Relative Chrolophyll content under 4ppm arsenic stress	0.548	0.010
Relative shoot weight under 6ppm arsenic stress Relative Chrolophyll content under 6ppm arsenic stress	0.073	0.754
Relative shoot weight under 4ppm arsenic stress Relative shoot weight under 6ppm arsenic stress	0.727	0.000
Relative Chrolophyll content under 4ppm arsenic stress Relative Chrolophyll content under 6ppm arsenic stress	0.027	0.907

5.3.9 At5G05560 was found from fine mapping and GWAS analysis

One locus located 1.6cM from the top of the consensus genetic map of chromosome 5, had an approximately additive effect on arsenite tolerance, explaining 21.8% of the genetic variation in this population. Fine mapping and GWAs narrowed the region. Within this region, there were 7 genes located. The presence of polymorphisms between the sequences of the parental accessions Ler-0 and Kas-2 was analyzed

using the GBrowse Viewers (http://signal.salk.edu/atg1001/3.0/gebrowser.php). Because Col-0 was used as the reference accession in this website, I analyzed all three accessions including examining the open reading frame of all well-annotated genes. Based on the LC50 result obtained from chapter 4 (Figure 4.16), Col-0 and Ler-0 were both arsenite tolerant types, whereas Kas-2 was the arsenite sensitive type. Only one gene among them has non-synonymous amino acid variation in the exon of Col-0, Ler-0 and Kas-2 genome. This gene is At5g05560, which encodes E3 ubiquitin ligase complex subunit. Information of all amino acid variation is given in Figure 5.19 and Non-synonymous amino acid variation is given in Table 5.8.

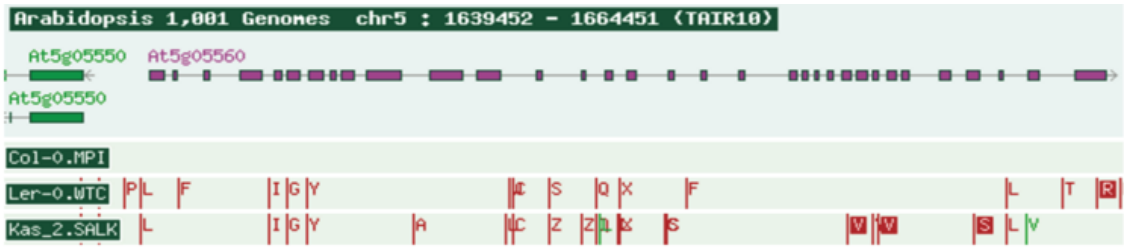

Figure 5.19. Allelic variation at At5G05560 involves structural polymorphisms. Information collected from SALK GBrowse Viewers.

Table 5.8 Non-synonymous amino acid variations between Col-0, Ler-0 and Kas-2 of chromosome 5 from 1623000 to 1658000 bp.

No.	Location on Chromosome 5	Col-0 As^{III} tolerance	Ler-0 As^{III} tolerance	Kas-2 As^{III} sensitive
1	1651109	Ile	Ile	Val
2	**1651531**	**Leu**	**Leu**	**Val**
3	1652854	Thr	Thr	Ser
4	1654568	Lys	Arg	Lys

The amino acid sequences were analyzed by using Clustal Omega web services from EMBL-EBI. The amino acid sequence alignment of Col-0, Ler-0 and Kas-2 was used to compare the amino acid difference. The second non-synonymous amino acid

variation is located in a conserved domain “Proteasome/cyclosome repeat; pfam01851” (CDD:280095).

The promoter sequence of gene At5G05560 was on database AtcisDB (TAIR9):

TAGAAAAGACTAGGGTTTAAGCGCTAAGACAACTTGGAAAGAGCGGGAGA

GTGTAAATTAAACTTCACCAATCTCCATTGCCTGTTGGCTTCGGAGAAAA

No difference of the promoter sequence was found between the three accessions Col-0, Ler-0 and Kas-2.

The transcriptome analysis under different experimental conditions were found on eFP Browser (http://bar.utoronto.ca/). The expression level of this gene was up-regulated under abiotic stress and the details are presented in Figure 5.20.

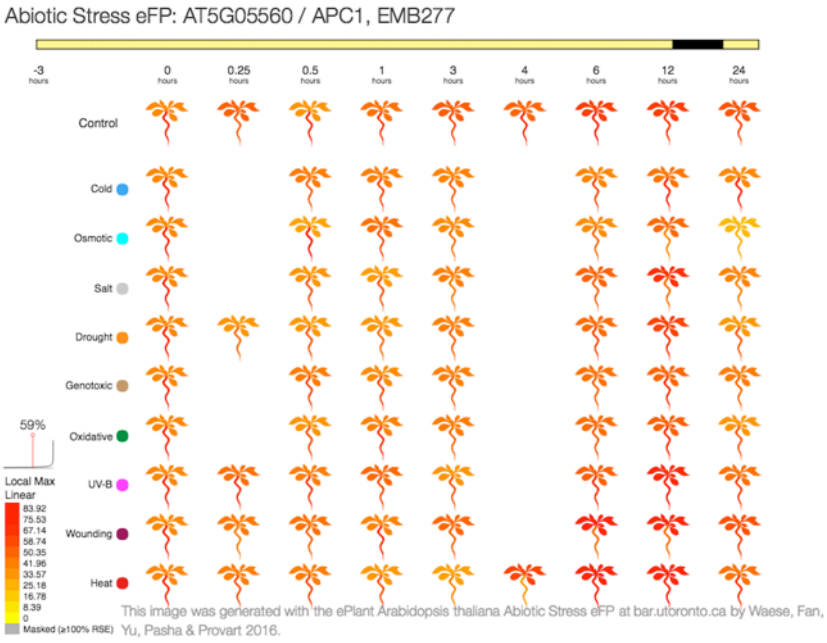

Figure 5.20. The transcriptome analysis of At5G05560 under abiotic stress. Published data obtained from the eFP database.

Twelve hours after the UV-B, wounding and heat stress had been given, the expression level of this gene was up regulated two times when compared with control. All current results suggest that this gene, which is located in the mapped region, is very likely to be an arsenite tolerance gene.

5.4 Discussion

5.4.1 Analysis method in QTL mapping

QTL mapping has been widely used in many plant researches. Many QTLs associated with valuable traits were identified using this approach. In the past, it has been used in finding naturally-occurring variation in flowering time of (Kowalski et al. 1994b). Recently, it was used to study salt tolerance in rice (*Oryza sativa*) seedlings under different salinity conditions (Wang et al. 2012). It was conducted to identify QTLs controlling the dimethylarsinic acid (DMA) content of rice (*Oryza sativa*) grains (Kuramata et al. 2013). And it was used in finding QTL underlying biomass yield and plant height in Switchgrass (Panicum virgatum) for bioenergy feedstock development (Serba et al. 2015).

Ten to fifteen years ago, the studies of QTL analysis were more likely to focus on theoretical research (Borevitz and Chory 2004; Collard et al. 2005; Price 2006). Most of the analysis methods were developed at that time. Three primary QTL mapping analysis methods were used in this research. The single marker analysis, also referred as single point analysis, is the simplest method to search for QTL associated with interesting trait using single marker. The statistical method used for the single point analysis includes T-test, analysis of variance (ANOVA) and linear regression (Ott and Hoh 2000). When using single marker analysis, each marker locus is independent of information for other loci, and analysis does not require complete linkage map. However, the major disadvantage of this analysis is that the further QTL is from a marker, the less likely it will be detected. Another method is interval mapping. This method makes use of linkage maps and analysis of intervals between adjacent pairs of linked markers along the chromosomes simultaneously (Ott and Hoh 2000). In a broad sense, the presence of a putative QTL is estimated if the logarithm of odds ratios (LOD) exceeds a critical threshold which is often fixed as > or =3 (Nyholt 2000). Interval mapping has been developed into composite interval mapping which is one of

the methods used to detect QTLs now. Composite interval mapping combines internal mapping with linear regression to analyze a marker interval with a few other well-chosen single markers in each analysis and provides more precise and effective predictions at mapping (Lou et al. 2006).

Segregation of resistance to arsenite was studied in the recombinant inbred population created from accessions Ler-0, the arsenite tolerant parent, and Kas-2, the arsenite sensitive parent. After measuring shoot weight of all RILs under 6ppm arsenite stress, the broad sense heritability (H^2) for both traits were estimated (for the method, see Cowling et al. 1997). Broad sense heritability for arsenite tolerance measurement was very high (absolute data: h^2=0.970: relative data: h^2=0.909) in the post-germination assay. This indicates that there is a strong genetic control of arsenite tolerance. The phenotypic measurement fitted into one normal distribution, which is a significant assumption required for these statistical methods that the phenotypic values of a trait can be modeled by a known parametric distribution (Wu et al. 2012).

The absolute measurement and relative measurement were two different traits and combining the both as the complex traits in the QTL analysis can improve the mapping accuracy. It was possible to get different genetic mapping results from using the absolute data and relative data, and indeed the absolute data analysis found more QTLs. Since absolute shoot weight was the result of both natural growth variation and the arsenite tolerance ability variation, the QTLs found in absolute data analysis would include both loci for nature growth and arsenite tolerance. In addition, the QTLs discovered from the relative data were consistent with some of the QTLs found from absolute data analysis, almost certainly, these were QTLs for arsenite tolerance.

The CIM analysis is more commonly used than IM analysis, but they were both used in this study because the results from the relative data analysis using the two methods suggest a minor difference in a major QTL. Analyzing two results together, QTL

mapping confirmed there were three QTLs on chromosome 5. By analyzing the absolute data, these three QTLs influence 50% of the phenotypic variance and 70% genetic variance. Using the relative data, these three QTL were accounted for 43% of the phenotypic variance and 87% genetic variance. The first one and second one were likely located close to marker NGA225 and SNP21 (Figure 5.1) and using LOD_2 data allowed us to target a region of 0mb~3.17mb on chromosome 5. The third one was likely located close to marker SNP358, M5-14 and SNP236 (Figure 5.1) and its physical region was in the region of 5.26mb~7.73mb on chromosome 5. In fine mapping analysis, 30 SNP markers covered these regions.

5.4.2 Using fine mapping refined the region of QTLs

The genotype data of 30 SNP markers of Ler-0 and Kas-2 from MALDI-TOF mass spectrometry-based SNP genotyping was compared with the known sequences to check the accuracy of the method. The result from it was as expected. After calculation, the genetic map showed a high-density linkage region between 0~1.3mb, and it increased the mapping difficulty. The fine mapping data supported the previous conclusion that three QTLs existed on chromosome 5. In the fine mapping analysis, the calculated threshold was lower than QTL mapping and this could be due to the smaller study region. The total genome size of is 135mb (Silady et al. 2011), and the size of the fine mapping is 7.3mb. The smaller the region is, the less significant the statistics data will be, but the observed peak in the fine mapping analysis was still higher than LOD3 in all analysis. Traditionally, a LOD score of 3 indicates that the odds that the loci that are linked are 1000 times greater than the odds that they are not linked (Nyholt 2000). Fine mapping analysis found three QTLs located on 27cM, 46cM and 82cM. In the fine mapping analysis, the regions where the first two QTLs located were no longer overlapped. The first two QTLs were located at 1.5mb~1.7mb and 5.2mb~5.5mb, respectively. The third QTL was at the end of the region, which was likely located in another high-density linkage region.

5.4.3 GWAS data analysis

The Figures show a genome-wide pattern of significance of association with arsenite tolerance traits: higher points represent more significantly associated SNPs. For the tolerance analysis, the best performing method seems to be GWAS with genetic principal components as cofactors. This method gave 4 out of the top 10 most significant SNP hits in the same region of chromosome 5 from 1623000bps to 1658000bps (personal communication with Dr. A. Brennan). Seven genes locate there and the gene At5G05560 was identified.

The tolerance and chlorophyll traits did not seem to detect the same SNPs or regions. The observation was also supported by the phenotypic character analysis. A significant correlation of 0.727 between shoot weight change under 4ppm and 6ppm arsenite stress was observed. This result was consistent with the previous conclusion in this study that the increased arsenite concentration could cause a decreased biomass accumulation. A weak correlation also existed between the shoot weight change and the chlorophyll content change under 4ppm stress. However, there was no link between the chlorophyll content changes from 4ppm stress to 6ppm stress, suggesting the chlorosis was not consistent with increased arsenite stress in tested *A thaliana* genotypes. The metabolism that works to regulate shoot weight (biomass) change under arsenite stress, did not help to preserve chlorophyll content. Combining with GWAS analysis, it can be concluded that shoot weight change and chlorophyll content change are two complex traits revealing different mechanisms.

A. thaliana, as a native species of the Eurasian continent, was introduced to other parts of the world, including North America, Japan, and East Australia through human activities (Silady et al. 2011). The 39 accessions used in GWAS analysis cover a wide geographical region including North America and Central Asia, but are mainly collected from Europe. After examining the geographical distribution effect on As tolerance variation, it seems the variation of As tolerance of the accessions used in

this research dose not come from natural selection, since Europe is not a high As-contaminated region. Therefore, it is possible that the mechanism of arsenite resistance may not be arsenite specific but more likely to act for general abiotic stress.

5.4.4 Analysis of gene At5G05560

Gene At5G05560 was found in one of the three QTLs. It encodes a subunit of the *Arabidopsis thaliana* E3 ubiquitin ligase complex. This gene has been reported that located in anaphase-promoting complex, nucleus, ubiquitin ligase complex, and expressed during four leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage (information was obtained from TAIR: http://www.arabidopsis.org/index.jsp). The function analysis of this gene was only reported in an anaphase-promoting complex/cyclosome (APC/C) study (Wang et al. 2013), and has never been reported for As^{III} resistance. As far as been known, At5g05560 corresponds to the largest APC/C subunit (APC1), which also shares some homology with two subunits of the 26S proteasome (Lupas et al. 1997). Regarding the studies of the E3 ubiquitin ligases, a functional genomics analysis pointed out that about 5% of *Arabidopsis* genome codes for proteins of ubiquitination pathway, and most of them (more than one thousand genes) correspond to E3 ubiquitin ligases that specifically recognise target proteins (Mazzucotelli et al. 2006). A huge number of members are in this family, which are involved in regulation of a number of biological processes including hormonal control of vegetative growth, plant reproduction, light response, biotic and abiotic stress tolerance and DNA repair, indicates a major role for protein degradation in control of plant life (Duplan and Rivas, 2014).

This gene was targeted in this gene mapping research. In our research, some allelic variation was found in this gene between Ler-0 and Kas-2, Moreover, the

non-synonymous amino acid difference in a conserved domain between Ler-0 and Kas-2 was found as well. Studies have also found it is up-regulated under the stresses caused by UV-B, wounding (punctured with pins) and 38°C (Kilian et al. 2007). Currently, no research has been done on the study of At5G05560 mutant or the phenotypic change when altering the gene expression level, which should be included in the future work of this study.

5.4.5 Summary

The final stage of this research is to employ Ler-0 x Kas-2 RIL mapping populations to find QTLs associated with plant As tolerance. Based on the QTL mapping result, a fine mapping targeting specific small regions of chromosome 5 in the *A. thaliana* genome was also accomplished. A small scale GWAS analysis was also used to integrate with fine mapping analysis to find As(III) tolerance-related genes. Homogeneous analysis suggested that the Ler-0 allele is quite dominant over the Kas-2 allele. Linking all the data obtained from genetic mapping, gene At5G05560 was found very likely to be an arsenite tolerance gene. However, without the functional analysis of this gene, it is difficult to draw any more conclusions. T-DNA insertional mutagenesis analysis and transcription analysis are highly recommended for the understanding of gene At5G05560.

Chapter 6

General Discussion and Future Plans

6.1 The importance of this study

This research was firstly aimed to identify plant ecotypes with enhanced arsenite tolerance. Hence, the physiological responses to arsenite stress were examined in two species of *Brassicaceae* family, the Indian mustard (*B. juncea*) and the model plant species, *A. thaliana*. Although plant As transport, metabolism and tolerance was an extensively studied research field, there were still many open questions regarding the genetic mechanisms of As tolerance, and so the outcomes from this work could be useful and important addition to the field. After studying the physiological responses to arsenite stress, this research employed QTL mapping and genome-wide association mapping to identify genetic mechanisms that control arsenite acquisition and tolerance. The potential gene targets were discovered through the progress, to provide benefits for future breeding or genetic manipulation strategies for generating 'safer' crop plants or for bioremediation applications. The novel contribution of this research included the developing of a physiological approach to identify the characteristics of arsenite tolerance on *B. juncea*. Moreover, a new *Arabidopsis* arsenite tolerance assays was invented. Under the arsenite stress, the physiological reaction of *B. juncea* and *A. thaliana* displayed distinct As accumulation and arsenite tolerance characteristics. In particular, the potential of future *B. juncea* breeding programmers and the potential to use *Arabidopsis* genetic resources for gene mapping have been this demonstrated. In addition, a candidate gene for arsenite tolerance was identified by genetic mapping, which has not been previously implicated in As tolerance. The aim of the candidate gene mapping was build upon previous assumptions from the literature and make use of previously developed *Arabidopsis* genetic resources. And the gene mapping approaches were used in an attempt to locate arsenite tolerance loci.

6.2 The selection of arsenite tolerance index for genetic mapping

There were two reasons why inorganic arsenite was chosen as the As stress species in all the experiments. The first reason is the high toxicity of arsenite, which is 60 times greater than arsenate (Hossain 2006). The second reason is that the metabolic reactions of plants towards arsenate are well established (Fitz and Wenzel 2002; Shin et al. 2004; Sánchez-Bermejo et al. 2014). Arsenate has drawn much public attention regarding food contamination resulting from its chemical similarity to inorganic phosphate (Pi), resulting in its easy transport across the plasmalemma by Pi transporter proteins (Finnegan and Chen 2012). As reviewed in the Introduction, the pathways of As in plants include the absorption of arsenate, arsenite and organic As by roots. The majority of As(V) and As (III) elements become As(III)-thiols, and they are sequestrated in the cell vacuoles. Whilst all arsenate is apparently reduced to arsenite in plant cells, the reason for the conversion of less toxic element As(V) to As(III) is still unknown. Similarly there is no clear plant arsenite detoxification mechanism. As little was known at the outset of the work presented in this thesis, the research focused on studying the damage caused by arsenite as the As species in all experiments in an attempt to improve the understanding of the plant arsenite tolerance mechanism.

Quantification of the phenotype associated with arsenite stress was crucial in this study. Genetic mapping was used in this study to find the linkage underlying the phenotype and the genotype. Currently, the genotyping approach and sequencing technology can provide accurate genotype data. However, consistent and accurate measurement of arsenite tolerance trait was formally difficult, because the behaviour of the plant under arsenite stress interferes with many other metabolic pathways as well as environmental effects. Moreover, the quantified arsenite tolerance phenotypes need to show the quantifiable difference between multiple accessions. In many abiotic stress studies, the relative growth rate (RGR) has been used. However, general plant

growth analysis was not suitable for this study because it required measuring the same trait at least twice at different time points (Hunt 1982). Traits like shoot weight cannot be measured twice, whilst root length and rosette diameter can be measured more than once without contaminating the experiment. However, it has been verified that root length and rosette are not suitable parameters in the present work. The data of root length contained large statistical deviation, whereas the difference in the relative rosette diameter between genotypes was not significant. The results from Chapter 4 demonstrated that the germination and post-germination stage needed testing separately. Thus, the shoot weight of 14-day old seedlings collected from the post-germination assay was used as the arsenite tolerance parameter.

Data collected from the post-germination assay were converted into two groups, the absolute phenotype measurement and the relative phenotype measurement. These two data sets were used as different arsenite tolerance index, and considered to be complex traits for QTL mapping and GWAS. The analysis for complex traits in multiple environments could increase the accuracy of QTL mapping (van Eeuwijk et al. 2010). The broad-sense heritability of absolute shoot weight and relative shoot weight, which were obtained from RIL mapping population were both very high - 97% for absolute data and 90.0% for relative data. This indicated that 97% of the variance in absolute shoot weight of the RIL population was caused by genetic influence. And 90.0% of the variance in relative shoot weight of the RIL population was caused by genetic influence. These very high broad-sense heritability data allowed using absolute data and relative data for QTL mapping. In the genetic mapping analysis, using absolute data resulted in a higher LOD value in QTL analysis. This is probably because more errors were introduced into the analysis using relative data. This kind of error is considered inevitable in growth rate data analysis.

From the plant arsenite tolerance phenotype measurements in *A. thaliana*, many other arsenite related parameters in addition to shoot weight were found to reveal different mechanisms related to arsenite for potential future genetic mapping. In

Chapter 3, As content in plant tissue was tested and plant arsenite transportability in xylem was discussed. The accession Nd-1 appears to have the highest transportability whereas genotype St-0 has the lowest transportability. The LC50 data of both Nd-1 and St-0 belonged to the middle As tolerance level, suggesting they had similar shoot weight change under arsenite stress. Theoretically, Nd-1 and St-0 can produce similar amount of shoot biomass, but the shoot of Nd-1 would contain more As than St-0 at a given applied As concentration. A mapping population from Nd-1 and St-0 could uncover the genetic and hence physiological basis for this differing arsenite xylem transportability in plant. Bur-0 had the highest As concentration in roots whereas Ler-0 had the lowest. A mapping population from Ler-0 and Bur-0 could therefore help understand the varied As root transport and As root detoxification.

The chlorophyll content of 39 genotypes was presented in Chapter 5. The etiolation in plant shoot and the reduction of the biomass under the same arsenite stress found by the GWAs analysis were caused by different As related loci and may underlie different genetic metabolisms. To discover genes or loci related to etiolation change under arsenite stress, any two or more genotypes with different chlorophyll content change caused by arsenite can be used to make a novel genetic mapping population to uncover the underlying genetic information.

6.3 Suggestion for future work

Genetic mapping is a highly effective approach for searching candidate genes associated with valuable traits. does not rely on the study of gene function. These approaches provides quantitative trait loci that belong to all types of ontological classes, not only the gene encoding enzymes of primary and secondary metabolism, metal transporters, but also the gene encoding transcription factors, signal transduction components, as well as genes with unknown functions (Silady et al. 2011).

Instead of generating a *de novo* mapping population, the mapping population employed in this research already existed, having been designed for other research purpose i.e the study of flowering time (El-Lithy et al. 2006). The biggest benefit of using a ready-made population is the time efficiency. Although it is technically simple to generate a RIL population in *A. thaliana* due to its short generation time and straightforward self-fertilisation, the process nevertheless requires up to two years to obtain a population. Furthermore, making the associated genetic map takes time and considerable financial resources. In some cases, where one of the parental lines is a late flowering accession, this progress can be even longer. Fortunately, the approach adopted in this work discovered a suitable mapping population (Ler-0/Kas-2) via testing arsenite tolerance variation among a large group of accessions.

It should be noted however, that the disadvantage of using a ready-made population is its fixed mapping resolution. To overcome this issue, it was decided to insert more markers into certain region of the genome for fine mapping (Rant et al. 2013). A total of 30 molecular markers (SNP marker) were added in the current mapping population to increase the mapping resolution. The accuracy of the designed SNP markers was confirmed to display the difference between accessions Ler-0 and Kas-2. An alternative method will be to generate a secondary mapping population like Near Isogenic Lines (NILs) population and this is suggested future studies. The 30 new SNP marks would be utilized again in the mapping of using of NILs mapping population.

By using the arsenite tolerance trait selection, arsenite tolerance mapping population selection, QTL mapping analysis, fine mapping analysis and GWAS analysis, I have discovered a gene At5G05560 that is strongly associated with arsenite/ heavy metal tolerance. Having found a gene accounting for shoot weight change under arsenite stress, among the remaining challenges is to demonstrate that the nucleotide polymorphisms are indeed involved in adaptation to arsenite environments. The real time PCR analysis is one approach to be considered here. If the expression of the

gene At5g05560 is different between the arsenite tolerance genotype and the arsenite sensitive genotype, this gene is likely to be involved in defending against arsenite stress. Complex gene networks control the tolerance responses of plants to many abiotic stresses; hence a transcript expression study in the presence of arsenite could help to build a framework and to understand the regulation in detail (Kilian et al. 2007).

Apart from gene At5g05560, more genes could be discovered in the target region of chromosome 5 of *A. thaliana,* since three loci with high additive effect were observed in QTL mapping. In future research, the next step can be, but not limited to, candidate gene selection (Pflieger et al. 2001). According to the annotated gene function, e.g. biotic and abiotic stress related proteins or analysis transcription factors and germin-like proteins, the gene related to arsenite tolerance can be selected. Thus, the function of the target gene under arsenite stress could be examined using mutation analysis, by using T-DNA insertion lines of the target genes, and/or the transgenic analysis that allows a researcher to detect the phenotype change caused by altering the gene expression level in plants.

As discussed in Chapter 5, accessions used in QTL mapping and GWAS analysis were not from a high As contamination site. Therefore, all these accessions may not have developed a particular metabolism for arsenite tolerance. Supporting evidence are observable phenotype changes such as anthocyanin accumulation under arsenite stress. The response is similar to general abiotic stress, for instance, following UV-B radiation. Previous research on maize found that As interfered with a set of maize root proteins including a functionally homogeneous group of 7 enzymes involved in cellular homeostasis for redox perturbation (e.g., superoxide dismutases, glutathione peroxidases, peroxiredoxin, and p-benzoquinone reductase), as well as four additional, functionally heterogeneous proteins (e.g., ATP synthase, succinyl-CoA synthetase, cytochrome P450 and guanine nucleotide-binding protein β subunit) (Requejo and Tena 2005). This finding suggests that the induction of oxidative stress

is the primary process underlying As toxicity in plants. As reviewed earlier, one molecular mechanism to decrease heavy metal toxicity in the plant is through auto-oxidation and fenton reaction to produce reactive oxygen species (Schützendübel and Polle 2002). To summarize, the metabolism of plant arsenite tolerance is more likely to have developed to combat general abiotic stress. Regarding the gene At5g05560 and other loci found from fine mapping analysis and GWAS analysis, and to test whether the major QTLs and polygenes found in this research are As specific or for the general heavy mental stress, QTLs analysis can be carried out using other toxic metals such as mercury, antimony, cadmium and lead (Schützendübel and Polle 2002). The post-germination assay can also be used for examining these heavy metals. The genotype analysis method used in As tolerance study can again be analyzed in other heavy metal studies

Comparative genetics among *A. thaliana* and other Brassicaceae species has been carried out for several other traits, such as seed colour and disease resistance (Schmidt et al. 2001). After establishing gene function of *A. thaliana*, homologous genes might also be found in *B. juncea* (Silady et al. 2011). Comparative mapping could be used to identify linkage of specific genes to QTLs of As tolerance-related traits between *A. thaliana* and *B. juncea*. The feasibility of this approach is supported by a study reported by Vatamaniuk et al. (1999), where the As tolerance gene *AtPCS1* from *A. thaliana* was used to identify the homologous gene *BjPCS1* in *B. juncea*. At the early stage of the research presented in this thesis, the As absorption of *B. juncea* was examined through hydroponic cultivation. This clearly demonstrated that seed oil harvested from a As contaminated environment would contain As and may not safe for consumption. By improving the understanding of *A. thaliana* arsenite tolerance, associated molecular markers could be employed to screen for tolerance-related genes in *Brassica* genotypes and to breed an ideal genotype for growing in As contaminated land.

References

Ahmad, P., Sarwat, M., Bhat, N. A., Wani, M. R., Kazi, A. G. and Tran, L. S. P., 2015. Alleviation of cadmium toxicity in *Brassica juncea* L. (Czern. & Coss.) by calcium application involves various physiological and biochemical strategies. *PloS One, **10***(1), p.e0114571.

Ahsan, N., Lee, D.G., Alam, I., Kim, P.J., Lee, J.J., Ahn, Y.O., Kwak, S.S., Lee, I.J., Bahk, J.D., Kang, K.Y. and Renaut, J., 2008. Comparative proteomic study of arsenic‐induced differentially expressed proteins in rice roots reveals glutathione plays a central role during As stress. *Proteomics, **8***(17), pp.3561-3576.

Arabidopsis eFP Browser: http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi

Ali, N. and Shah, S.A.H., 1982. Seed yield and oil contents in *Poorbi* Raya (*Brassica juncea*) as affected by number and time of irrigations. *Pakistan Journal of Agricultural Research, **3***(2), pp.138-140.

Åkesson, M., Bensch, S., Hasselquist, D., Tarka, M. and Hansson, B., 2008. Estimating heritabilities and genetic correlations: comparing the 'animal model' with parent-offspring regression using data from a natural population. *PLoS One, **3***(3), p.e1739.

Alonso-Blanco, C., Mendez-Vigo, B. and Koornneef, M., 2005. From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. *International Journal of Developmental Biology, **49***(5-6), pp.717-732.

Asimit, J. and Zeggini, E., 2010. Rare variant association analysis methods for complex traits. *Annual Review of Genetics, **44***, pp.293-308.

Atkins, P.J., Hassan, M. and Dunn, C.E., 2006. Toxic torts: arsenic poisoning in Bangladesh and the legal geographies of responsibility. *Transactions of the Institute of British Geographers, **31***(3), pp.272-285.

Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A.M., Hu, T.T. and Jiang, R., 2010. Genome-wide

association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature, **465***(7298), pp.627-631.

Azcue, J.M. and Nriagu, J.O., 1994. Arsenic. historical perspectives. *Advances in Environmental Science and Technology, **26***, pp.31-32.

Baker, A.J.M., McGrath, S.P., Reeves, R.D., Smith, J.A.C., Terry, N. and Bañuelos, G., 2000. Phytoremediation of contaminated soil and water. *Phytoremediation of Contaminated Soil and Water*, pp.2231-2232.

Barnes, S., 2002. Comparing *Arabidopsis* to other flowering plants. *Current Opinion in Plant Biology, **5***(2), pp.128-134.

Barrachina, A.C., Carbonell, F.B. and Beneyto, J.M., 1995. Arsenic uptake, distribution, and accumulation in tomato plants: effect of arsenite on plant growth and yield. *Journal of Plant Nutrition*, ***18***(6), pp.1237-1250.

Beadle, C.L., 2014. Plant growth analysis. *Techniques in Bioproductivity and Photosynthesis, **2**,* pp.20-25.

Beemster, G.T., De Vusser, K., De Tavernier, E., De Bock, K. and Inzé, D., 2002. Variation in growth rate between *Arabidopsis* ecotypes is correlated with cell division and A-type cyclin-dependent kinase activity. *Plant Physiology, **129***(2), pp.854-864.

Bhumbla, D.K., and Keefer, R.F., 1994. Arsenic mobilization and bioavailability in soils. *Arsenic in the Environment: part 1. Cycling and Characterization*, pp.51-82.

Blankendaal, M., 1972. Growing plants without soil for experimental use (No. 1251). Agricultural Research Service, US Department of Agriculture.

Blaylock, M.J., Salt, D.E., Dushenkov, S., Zakharova, O., Gussman, C., Kapulnik, Y., Ensley, B.D. and Raskin, I., 1997. Enhanced accumulation of Pb in Indian mustard by soil-applied chelating agents. *Environmental Science and Technology*, ***31***(3), pp.860-865.

Blievernicht, J.K., Schaeffeler, E. and Klein, K., 2007. MALDI-TOF mass spectrometry for multiplex genotyping of CYP2B6 single-nucleotide polymorphisms. *Clinical Chemistry, **53***(1), pp.24-33.

Borevitz, J.O. and Nordborg, M., 2003. The impact of genomics on the study of natural variation in *Arabidopsis*. *Plant Physiology, **132***(2), pp.718-725.

Brachi, B., Morris, G.P. and Borevitz, J.O., 2011. Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology,**12*(10), pp.97-98.

Bucher, M., 2007. Functional biology of plant phosphate uptake at root and mycorrhiza interfaces. *New Phytologist, **173***(1), pp.11-26.

Borevitz, J.O. and Chory, J., 2004. Genomics tools for QTL analysis and gene discovery. *Current Opinion in Plant Biology*, ***7***(2), pp.132-136.

Castillo-Michel, H., Parsons, J.G., Peralta-Videa, J.R., Martinez-Martinez, A., Dokken, K.M. and Gardea-Torresdey, J.L., 2007. Use of X-ray absorption spectroscopy and biochemical techniques to characterize arsenic uptake and reduction in pea (*Pisum sativum*) plants. *Plant Physiology and Biochemistry, **45***(6), pp.457-463.

Catarecha, P., Segura, M.D., Franco-Zorrilla, J.M., García-Ponce, B., Lanza, M., Solano, R., Paz-Ares, J. and Leyva, A., 2007. A mutant of the *Arabidopsis* phosphate transporter PHT1; 1 displays enhanced arsenic accumulation. *The Plant Cell, **19***(3), pp.1123-1133.

Chalker‐**Scott, L.**, 1999. Environmental significance of anthocyanins in plant stress responses. *Photochemistry and Photobiology, **70***(1), pp.1-9.

Chibuike, G. U., and Obiora, S. C. 2014. Heavy metal polluted soils: effect on plants and bioremediation methods. *Applied and Environmental Soil Science*, ***2014***, pp.112-124.

Chowdhury, U.K., Biswas, B.K., Chowdhury, T.R., Samanta, G., Mandal, B.K., Basu, G.C., Chanda, C.R., Lodh, D., Saha, K.C., Mukherjee, S.K. and Roy, S., 2000. Groundwater arsenic contamination in Bangladesh and West Bengal, India. *Environmental Health Perspectives, **108***(5), p.393.

Cobbett, C.S., 2000. Phytochelatins and their roles in heavy metal detoxification. *Plant Physiology, **123***(3), pp.825-832.

Cobbett, C. and Goldsbrough, P., 2002. Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. *Annual Review of Plant Biology, **53***(1), pp.159-182.

Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B. and Pang, E.C.K., 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica, **142***(1-2), pp.169-196.

Cowling, W.A., Sweetingham, M.W., Diepeveen, D. and Cullis, B.R., 1997. Heritability of resistance to brown spot and root rot of narrow-leafed lupins caused by Pleiochaeta setosa (Kirchn.) Hughes in field experiments. *Plant Breeding, **116***(4), pp.341-345.

Curry, A., Carrin, G., Bartram, J., Yamamura, S., Heijnen, H., Sims, J., Hueb, J. and Sato, Y., 2000. Towards an assessment of the socioeconomic impact of arsenic poisoning in Bangladesh. *World Health Organization*.

Dhankher, O.P., Rosen, B.P., McKinney, E.C. and Meagher, R.B., 2006. Hyperaccumulation of arsenic in the shoots of *Arabidopsis* silenced for arsenate reductase (ACR2). *Proceedings of the National Academy of Sciences,* U.S.A. ***103***(14), pp.5413-5418.

Do, L.D., Stevens, T.L., Kibbey, T.C. and Sabatini, D.A., 2014. Preliminary formulation development for aqueous surfactant-based soybean oil extraction. *Industrial Crops and Products, **62***, pp.140-146.

Doganlar, S., Frary, A. and Tanksley, S.D., 2000. The genetic basis of seed-weight variation: tomato as a model system. *Theoretical and Applied Genetics, **100***(8), pp.1267-1273.

Duplan, V. and Rivas, S., 2014. E3 ubiquitin-ligases and their target proteins during the regulation of plant innate immunity. *Frontiers in Plant Science*. ***5***(42). pp.1-6.

El-Lithy, M.E., Bentsink, L., Hanhart, C.J., Ruys, G.J., Rovito, D., Broekhof, J.L., van der Poel, H.J., Van Eijk, M.J., Vreugdenhil, D. and Koornneef, M., 2006. New *Arabidopsis* recombinant inbred line populations genotyped using SNPWave and their use for mapping flowering-time quantitative trait loci. *Genetics, **172***(3), pp.1867-1876.

El-Soda, M., Malosetti, M., Zwaan, B.J., Koornneef, M. and Aarts, M.G., 2014. Genotype× environment interaction QTL mapping in plants: lessons from *Arabidopsis*. *Trends in Plant Science, **19***(6), pp.390-398.

Ercal, N., Gurer-Orhan, H. and Aykin-Burns, N., 2001. Toxic metals and oxidative stress part I: mechanisms involved in metal-induced oxidative damage. *Current Topics in Medicinal Chemistry, **1***(6), pp.529-539.

Erickson, M.L. and Barnes, R.J., 2003. Measured and Modeled Arsenic Species Variability in Midwestern Ground Water. In 2003 Geological Society of America Annual Meeting and Exposition.

Evans, G.C., 1972. The quantitative analysis of plant growth (Vol. 1). California: University of California Press.

Finnegan, P. and Chen, W. (2012). Arsenic toxicity: the effects on plant metabolism. *Frontiers in Physiology, **3***, p.182.

Fitz, W.J. and Wenzel, W.W., 2002. Arsenic transformations in the soil-rhizosphere-plant system: fundamentals and potential application to phytoremediation. *Journal of Biotechnology, **99***(3), pp.259-278.

Fourcaud, T., Zhang, X., Stokes, A., Lambers, H. and Körner, C., 2008. Plant growth modelling and applications: the increasing importance of plant architecture in growth models. *Annals of Botany, **101***(8), pp.1053-1063.

Fournier-Level, A., Korte, A., Cooper, M.D., Nordborg, M., Schmitt, J. and Wilczek, A.M., 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science, **334***(6052), pp.86-89.

Francis, F.J., 1982. *Analysis of anthocyanins*. Anthocyanins as Food Colors. American: Academic press, pp.181-207.

Frankenberger Jr, W.T. and Arshad, M., 2002. Volatilization of arsenic. *Environmental Chemistry of Arsenic*, pp.363-380.

Fu, Z., Li, W., Xing, X., Xu, M., Liu, X., Li, H., Xue, Y., Liu, Z. and Tang, J., 2016. Genetic analysis of arsenic accumulation in maize using QTL mapping. *Scientific Reports, **6***. pp.333.

Gasic, K. and Korban, S.S., 2007. Transgenic Indian mustard (*Brassica juncea*) plants expressing an *Arabidopsis* phytochelatin synthase (AtPCS1) exhibit enhanced As and Cd tolerance. *Plant Molecular Biology, **64**(*4), pp.361-369.

Ghanthimathi, S., Aminah, A., SAIMIJA, S., Ujang, T. and Nurul Izzah, A., 2012. Comparison of microwave assisted acid digestion methods for ICP-MS determination of total arsenic in fish tissue. *Sains Malaysiana, **41***(12), pp.1557-1564.

Ghosh, M., Shen, J. and Rosen, B.P., 1999. Pathways of As (III) detoxification in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences, **96***(9), pp.5001-5006.

Gómez-Ariza, J. L., Sánchez-Rodas, D., Giráldez, I. and Morales, E., 2000. A comparison between ICP-MS and AFS detection for arsenic speciation in environmental samples. *Talanta, **51***(2), pp.257-268.

Gould, K.S., McKelvie, J. and Markham, K.R., 2002. Do anthocyanins function as antioxidants in leaves? Imaging of H_2O_2 in red and green leaves after mechanical injury. *Plant, Cell & Environment, **25***(10), pp.1261-1269.

Grant, D. and Shoemaker, R.C., 2009. Plant gene mapping techniques. In: eLS. John Wiley & Sons Ltd, Chichester.

Ha, S.B., Smith, A.P., Howden, R., Dietrich, W.M., Bugg, S., O'Connell, M.J., Goldsbrough, P.B. and Cobbett, C.S., 1999. Phytochelatin synthase genes from *Arabidopsis* and the yeast *Schizosaccharomyces pombe*. *The Plant Cell, **11***(6), pp.1153-1163.

Heiss, S., Wachter, A., Bogs, J., Cobbett, C. and Rausch, T., 2003. Phytochelatin synthase (PCS) protein is induced in *Brassica juncea* leaves after prolonged Cd exposure. *Journal of Experimental Botany, **54***(389), pp.1833-1839.

Hemingway, J.S., 1976. Mustards: *Brassica* spp. and *Sinapis alba* (Cruciferae). *Evolution of Crop Plants*. pp.56-69.

Hoagland, D.R. and Arnon, D.I., 1950. The water-culture method for growing plants without soil. *Circular, **347***. Berkeley: The college of Agricultural, University of California.

Horiguchi, G., Ferjani, A., Fujikura, U. and Tsukaya, H., 2006. Coordination of cell proliferation and cell expansion in the control of leaf size in *Arabidopsis thaliana*. *Journal of Plant Research, **119***(1), pp.37-42.

Horneman, A., van Geen, A., Kent, D.V., Mathe, P.E., Zheng, Y., Dhar, R.K., O'connell, S., Hoque, M.A., Aziz, Z., Shamsudduha, M. and Seddique, A.A., 2004. Decoupling of As and Fe release to Bangladesh groundwater under reducing conditions. Part I: Evidence from sediment profiles. *Geochimica et Cosmochimica Acta, **68***(17), pp.3459-3473.

Hossain, M.F., 2006. Arsenic contamination in Bangladesh -- an overview. *Agriculture, Ecosystems & Environment, **113***(1), pp.1-16.

Hughes, M.F., 2002. Arsenic toxicity and potential mechanisms of action. *Toxicology Letters, **133***(1), pp.1-16.

Hunt, R., 1982. *Plant growth analysis*. London: Institute of Terrestrial Ecology.

Hunt, R., Thomas, B., Murphy, D.J. and Murray, D., 2003. Growth analysis, individual plants. *Encyclopedia of Applied Plant Sciences, **2***, pp.579-588.

Huq, S., Joardar, J. C., Parvin, S. and Correll, R., 2006. Arsenic contamination in food-chain: transfer of arsenic into food materials through groundwater irrigation. *Journal of Health, Population and Nutrition*, pp.305-316.

Irvine, L., Boyer, I.J. and DeSesso, J.M., 2006. Monomethylarsonic acid and dimethylarsinic acid: developmental toxicity studies with risk assessment. *Birth Defects Research Part B: Developmental and Reproductive Toxicology, **77***(1), pp.53-68.

Isayenkov, S.V. and Maathuis, F.J., 2008. The *Arabidopsis thaliana* aquaglyceroporin AtNIP7; 1 is a pathway for arsenite uptake. *Febs Letters, **582***(11), pp.1625-1628.

Jackson, B.P. and Punshon, T., 2015. Recent Advances in the Measurement of Arsenic, Cadmium, and Mercury in Rice and Other Foods. *Current Environmental Health Reports, **2***(1), pp.15-24.

Järup, L., 2003. Hazards of heavy metal contamination. *British Medical Bulletin, **68***(1), pp.167-182.

John, R., Ahmad, P., Gadgil, K. and Sharma, S., 2012. Heavy metal toxicity: Effect on plant growth, biochemical parameters and metal accumulation by *Brassica juncea L*. *International Journal of Plant Production,* ***3***(3), pp.65-76.

Jorde, L. B., 2000. Linkage disequilibrium and the search for complex disease genes. *Genome Research,* ***10***(10), pp.1435-1444.

Jurinke, C., Oeth, P. and van den Boom, D., 2004. MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis. *Molecular Biotechnology,* ***26***(2), pp.147-64.

Juszczuk, I.M., Wiktorowska, A., Malusá, E. and Rychter, A.M., 2004. Changes in the concentration of phenolic compounds and exudation induced by phosphate deficiency in bean plants (*Phaseolus vulgaris* L.). *Plant and Soil,* ***267***(1-2), pp.41-49.

Kang, Y.H., Yi, M.J., Kim, M.J., Park, M.T., Bae, S., Kang, C.M., Cho, C.K., Park, I.C., Park, M.J., Rhee, C.H. and Hong, S.I., 2004. Caspase-Independent Cell Death by Arsenic Trioxide in Human Cervical Cancer Cells Reactive Oxygen Species-Mediated Poly (ADP-ribose) Polymerase-1 Activation Signals Apoptosis-Inducing Factor Release from Mitochondria. *Cancer Research,* ***64***(24), pp.8960-8967.

Karimi, N., Ghaderian, S.M., Raab, A., Feldmann, J. and Meharg, A.A., 2009. An arsenic‐accumulating, hypertolerant brassica, Isatis capadocica. *New Phytologist,* ***184***(1), pp.41-47.

Kaul, S., Koo, H.L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L.J., Feldblyum, T., Nierman, W., Benito, M.I., Lin, X. and Town, C.D., 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature,* ***408***(6814), pp.796-815.

Kearsey, M.J. and Farquhar, A.G.L., 1998. QTL analysis in plants; where are we now? *Heredity,* ***80***(2), pp.137-142.

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., D'Angelo, C., Bornberg‐Bauer, E., Kudla, J. and Harter, K., 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV‐B light, drought and cold stress responses. *The Plant Journal,* ***50***(2), pp.347-363.

Korte, A. and Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods,* ***9***(1), p.1.

Korte, A., Vilhjálmsson, B.J., Segura, V., Platt, A., Long, Q. and Nordborg, M., 2012. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics,* ***44***(9), pp.1066-1071.

Kover, P.X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I.M., Purugganan, M.D., Durrant, C. and Mott, R., 2009. A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genetic,* ***5***(7), p.e1000551.

Kovinich, N., Kayanja, G., Chanoca, A., Otegui, M.S. and Grotewold, E., 2015. Abiotic stresses induce different localizations of anthocyanins in *Arabidopsis*. *Plant Signaling & Behaviour,* ***10***(7), p.e1027850.

Kowalski, S.P., Lan, T.H., Feldmann, K.A. and Paterson, A.H., 1994a. Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics,* ***138***(2), pp.499-510.

Kowalski, S.P., Lan, T.H., Feldmann, K.A. and Paterson, A.H., 1994b. QTL mapping of naturally-occurring variation in flowering time of *Arabidopsis thaliana*. *Molecular and General Genetics* MGG, ***245***(5), pp.548-555.

Kroymann, J. and Mitchell-Olds, T., 2005. Epistasis and balanced polymorphism influencing complex trait variation. *Nature,* ***435***(7038), pp.95-98.

Kuramata, M., Abe, T., Kawasaki, A., Ebana, K., Shibaya, T., Yano, M. and Ishikawa, S., 2013. Genetic diversity of arsenic accumulation in rice and QTL analysis of methylated arsenic in rice grains. *Rice*, ***6***(1), pp.97-98.

Kumar, M., Choi, J.Y., Kumari, N., Pareek, A. and Kim, S.R., 2015. Molecular breeding in Brassica for salt tolerance: importance of microsatellite (SSR) markers for molecular breeding in Brassica. *Frontiers in Plant Science,* ***6***. pp.47-49.

Leão, G.A., de Oliveira, J.A., Felipe, R.T.A., Farnese, F.S. and Gusman, G.S., 2014. Anthocyanins, thiols, and antioxidant scavenging enzymes are involved in Lemna gibba tolerance to arsenic. *Journal of Plant Interactions,* ***9***(1), pp.143-151.

Lee, P.C., Kakadiya, R., Su, T.L. and Lee, T.C., 2010. Combination of bifunctional alkylating agent and arsenic trioxide synergistically suppresses the growth of drug-resistant tumor cells. *Neoplasia,* ***12***(5), p.376.

Lee, J., Rennaker, C. and Wrolstad, R.E., 2008. Correlation of two anthocyanin quantification methods: HPLC and spectrophotometric methods. *Food Chemistry,* ***110***(3), pp.782-786.

Levine, J.M., Adler, P.B. and Yelenik, S.G., 2004. A meta‐analysis of biotic resistance to exotic plant invasions. *Ecology Letters,* ***7***(10), pp.975-989.

Li, R.Y., Ago, Y., Liu, W.J., Mitani, N., Feldmann, J., McGrath, S.P., Ma, J.F. and Zhao, F.J., 2009. The rice aquaporin Lsi1 mediates uptake of methylated arsenic species. *Plant Physiology,* ***150*** (4), pp.2071-2080.

Li, Z.S., Lu, Y.P., Zhen, R.G., Szczypka, M., Thiele, D.J. and Rea, P.A., 1997. A new pathway for vacuolar cadmium sequestration in *Saccharomyces cerevisiae*: YCF1-catalyzed transport of bis (glutathionato) cadmium. *Proceedings of the National Academy of Sciences,* ***94***(1), pp.42-47.

Liu, C., Weng, J., Zhang, D., Zhang, X., Yang, X., Shi, L., Meng, Q., Yuan, J., Guo, X., Hao, Z. and Xie, C., 2014. Genome-wide association study of resistance to rough dwarf disease in maize. *European Journal of Plant Pathology,* ***139***(1), pp.205-216.

Liu, J., Huang, S., Sun, M., Liu, S., Liu, Y., Wang, W., Zhang, X., Wang, H. and Hua, W., 2012. An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods,* ***8***(1), p.1.

Lou, X.Y., Ma, J.Z., Yang, M.C., Zhu, J., Liu, P.Y., Deng, H.W., Elston, R.C. and Li, M.D., 2006. Improvement of mapping accuracy by unifying linkage and association analysis. *Genetics,* ***172***(1), pp.647-661.

Lupas, A., Baumeister, W. and Hofmann, K., 1997. A repetitive sequence in subunits of the 26S proteasome and 20S cyclosome (anaphase-promoting complex). *Trends in Biochemical Sciences*, ***22***(6), pp.195-196.

Ma, J.F., Tamai, K., Yamaji, N., Mitani, N., Konishi, S., Katsuhara, M., Ishiguro, M., Murata, Y. and Yano, M., 2006. A silicon transporter in rice. *Nature*, *440*(7084), pp.688-691.

Mahimairaja, S., Bolan, N.S., Adriano, D.C. and Robinson, B., 2005. Arsenic contamination and its risk management in complex environmental settings. *Advances in Agronomy, **86***, pp.1-82.

Martinez, V.D., Vucic, E.A., Becker-Santos, D.D., Gil, L. and Lam, W.L., 2011. Arsenic exposure and the induction of human cancers. *Journal of Toxicology, **2011.*** p.1.

Martinoia, E., Meyer, S., De Angeli, A. and Nagy, R., 2012. Vacuolar transporters in their physiological context. *Plant Biology, **63***(1), p.183.

Mazzucotelli, E., Belloni, S., Marone, D., De Leonardis, A.M., Guerra, D., Di Fonzo, N., Cattivelli, L. and Mastrangelo, A.M., 2006. The E3 ubiquitin ligase gene family in plants: regulation by degradation. *Current Genomics*, ***7***(8), pp.509-522.

Meharg, A.A. and Jardine, L., 2003. Arsenite transport into paddy rice (*Oryza sativa*) roots. *New Phytologist, **157***(1), pp.39-44.

Meharg, A.A., Naylor, J. and Macnair, M.R., 1994. Phosphorus nutrition of arsenate-tolerant and nontolerant phenotypes of velvetgrass. *Journal of Environmental Quality, 23*(2), pp.234-238.

Meharg, A.A. and Rahman, M.M., 2003. Arsenic contamination of Bangladesh paddy field soils: implications for rice contribution to arsenic consumption. *Environmental Science and Technology, **37***(2), pp.229-234.

Mendoza-Cózatl, D.G., Jobe, T.O., Hauser, F. and Schroeder, J.I., 2011. Long-distance transport, vacuolar sequestration, tolerance, and transcriptional responses induced by cadmium and arsenic. *Current Opinion in Plant Biology, **14***(5), pp.554-562.

Montefiori, M., Espley, R.V., Stevenson, D., Cooney, J., Datson, P.M., Saiz, A., Atkinson, R.G., Hellens, R.P. and Allan, A.C., 2011. Identification and characterisation of F3GT1 and F3GGT1, two glycosyltransferases responsible for anthocyanin biosynthesis in red‐fleshed kiwifruit (*Actinidia chinensis*). *The Plant Journal, **65***(1), pp.106-118.

Montes-Bayón, M., Meija, J., LeDuc, D. L., Terry, N., Caruso, J. A. and Sanz-Medel, A., 2004. HPLC-ICP-MS and ESI-Q-TOF analysis of biomolecules

induced in *Brassica juncea* during arsenic accumulation. *Journal of Analytical Atomic Spectrometry,* ***19***(1), pp.153-158.

Moreno-Jiménez, E., Esteban, E. and Peñalosa, J.M., 2012. The fate of arsenic in soil-plant systems. In Reviews of environmental contamination and toxicology (pp. 1-37). New York: Springer.

Moreno-Jiménez, E., Esteban, E., Fresno, T., de Egea, C.L. and Peñalosa, J.M., 2010. Hydroponics as a valid tool to assess arsenic availability in mine soils. *Chemosphere,* ***79***(5), pp.513-517.

Mouchel, C.F., Briggs, G.C. and Hardtke, C.S., 2004. Natural genetic variation in *Arabidopsis* identifies BREVIS RADIX, a novel regulator of cell proliferation and elongation in the root. *Genes and Development,* ***18***(6), pp.700-714.

Nakabayashi, R., Yonekura‐Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., Matsuda, F., Kojima, M., Sakakibara, H., Shinozaki, K. and Michael, A.J., 2014. Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. *The Plant Journal,* ***77***(3), pp.367-379.

Nyholt, D.R., 2000. All LODs are not created equal. *The American Journal of Human Genetics,* ***67***(2), pp.282-288.

Oraguzie, N.C., Gardiner, S.E., Rikkerink, E.H. and Silva, H.N. eds., 2007.Association mapping in plants (pp. 11-39). New York: Springer.

Ott, J. and Hoh, J., 2000. Statistical approaches to gene mapping. *The American Journal of Human Genetics,* ***67***(2), pp.289-294.

Päivöke, A.E. and Simola, L.K., 2001. Arsenate toxicity to *Pisum sativum*: mineral nutrients, chlorophyll content, and phytase activity. *Ecotoxicology and Environmental Safety,* ***49***(2), pp.111-121.

Patra, M., Bhowmik, N., Bandopadhyay, B. and Sharma, A., 2004. Comparison of mercury, lead and arsenic with respect to genotoxic effects on plant systems and the development of genetic tolerance. *Environmental and Experimental Botany,* ***52***(3), pp.199-223.

Peer, W.A., Brown, D.E., Tague, B.W., Muday, G.K., Taiz, L. and Murphy, A.S., 2001. Flavonoid accumulation patterns of transparent testa mutants of *Arabidopsis*. *Plant Physiology,* ***126***(2), pp.536-548.

Pérez-Pérez, J.M., Serrano-Cartagena, J. and Micol, J.L., 2002. Genetic analysis of natural variations in the architecture of *Arabidopsis thaliana* vegetative leaves. *Genetics,* ***162***(2), pp.893-915.

Pflieger, S., Lefebvre, V. and Causse, M., 2001. The candidate gene approach in plant genetics: a review. *Molecular Breeding,* ***7***(4), pp.275-291.

Pickering, I.J., Prince, R.C., George, M.J., Smith, R.D., George, G.N. and Salt, D.E., 2000. Reduction and coordination of arsenic in Indian mustard. *Plant Physiology,* ***122***(4), pp.1171-1178.

Platt, A., Vilhjálmsson, B.J. and Nordborg, M., 2010. Conditions under which genome-wide association studies will be positively misleading. *Genetics,* ***186***(3), pp.1045-1052.

Poorter, H., 1989. Plant growth analysis: towards a synthesis of the classical and the functional approach. *Physiologia Plantarum,* ***75***(2), pp.237-244.

Porra, R.J., 2005. The chequered history of the development and use of simultaneous equations for the accurate determination of chlorophylls a and b. In Discoveries in Photosynthesis (pp. 633-640). Netherlands: Springer.

Pourcel, L., Irani, N.G., Koo, A.J., Bohorquez‐Restrepo, A., Howe, G.A. and Grotewold, E., 2013. A chemical complementation approach reveals genes and interactions of flavonoids with other pathways. *The Plant Journal,* ***74***(3), pp.383-397.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A., 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding,* ***2***(3), pp.225-238.

Prakash, S., Bhat, S.R., Quiros, C.F., Kirti, P.B. and Chopra, V.L., 2009. *Brassica* and its close allies: cytogenetics and evolution. *Plant Breeding Reviews,* ***31****,* p.21.

Price, A.H., 2006. Believe it or not, QTLs are accurate! *Trends in Plant Science, **11***(5), pp.213-216.

Pusch, W., Wurmbach, J. H., Thiele, H. and Kostrzewa, M., 2002. MALDI-TOF mass spectrometry-based SNP genotyping. *Pharmacogenomics, **3***(4), pp.537-548.

Raab, A., Schat, H., Meharg, A.A. and Feldmann, J., 2005. Uptake, translocation and transformation of arsenate and arsenite in sunflower (*Helianthus annuus*): formation of arsenic-phytochelatin complexes during exposure to high arsenic concentrations. *New Phytologist, **168***(3), pp.551-558.

Rafalski, J.A., 2002. Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Science, **162***(3), pp.329-333.

Rahman, M., Haq, N. and Williams, I.D., 2012. Genetic effect on phytoaccumulation of arsenic in Brassica juncea L. *Euphytica, **186***(2), pp.409-417.

Rahman, M.A., Hasegawa, H., Rahman, M.M., Islam, M.N., Miah, M.M. and Tasmen, A., 2007. Effect of arsenic on photosynthesis, growth and yield of five widely cultivated rice (*Oryza sativa* L.) varieties in Bangladesh. *Chemosphere, **67***(6), pp.1072-1079.

Rant, J. C., Arraiano, L. S., Chabannes, M. and Brown, J. K., 2013. Quantitative trait loci for partial resistance to Pseudomonas syringae pv. maculicola in *Arabidopsis thaliana*. *Molecular Plant Pathology, **14***(8), pp.828-837.

Requejo, R. and Tena, M., 2005. Proteome analysis of maize roots reveals that oxidative stress is a main contributing factor to plant arsenic toxicity. *Phytochemistry, **66***(13), pp.1519-1528.

Rieseberg, L.H., Archer, M.A. and Wayne, R.K., 1999. Transgressive segregation, adaptation and speciation. *Heredity, **83***(4), pp.363-372.

Ritchie, R.J., 2006. Consistent sets of spectrophotometric chlorophyll equations for acetone, methanol and ethanol solvents. *Photosynthesis Research, **89***(1), pp.27-41.

Roeder, K., Bacanu, S.A., Sonpar, V., Zhang, X. and Devlin, B., 2005. Analysis of single-locus tests to detect gene/disease associations. *Genetic Epidemiology,* ***28***(3), pp.207-219.

Rofkar, J.R., Dwyer, D.F. and Bobak, D.M., 2014. Uptake and Toxicity of Arsenic, Copper, and Silicon in *Azolla caroliniana* and *Lemna minor*. *International Journal of Phytoremediation,* ***16***(2), pp.155-166.

Rosen, B.P., Ajees, A.A. and McDermott, T.R., 2011. Life and death with arsenic. *Bioessays,* ***33***(5), pp.350-357.

Rosenthal, A., Pyle, D.L. and Niranjan, K., 1996. Aqueous and enzymatic processes for edible oil extraction. *Enzyme and Microbial Technology,* ***19***(6), pp.402-420.

Salvi, S. and Tuberosa, R., 2005. To clone or not to clone plant QTLs: present and future challenges. *Trends in Plant Science,* ***10***(6), pp.297-304.

Sánchez-Bermejo, E., Castrillo, G., Del Llano, B., Navarro, C., Zarco-Fernández, S., Martinez-Herrera, D.J., Leo-del Puerto, Y., Muñoz, R., Cámara, C., Paz-Ares, J. and Alonso-Blanco, C., 2014. Natural variation in arsenate tolerance identifies an arsenate reductase in *Arabidopsis thaliana*. *Nature Communications,* ***5***. p.1.

Schat, H., Llugany, M., Vooijs, R., Hartley-Whitaker, J. and Bleeker, P.M., 2002. The role of phytochelatins in constitutive and adaptive heavy metal tolerances in hyperaccumulator and non-hyperaccumulator metallophytes. *Journal of Experimental Botany,* ***53***(379), pp.2381-2392.

Schmidt, R., Acarkan, A. and Boivin, K., 2001. Comparative structural genomics in the Brassicaceae family. *Plant Physiology and Biochemistry,* ***39***(3), pp.253-262.

Schmöger, M.E., Oven, M. and Grill, E., 2000. Detoxification of arsenic by phytochelatins in plants. *Plant Physiology,* ***122***(3), pp.793-802.

Schützendübel, A., and Polle, A., 2002. Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization. *Journal of Experimental Botany,* ***53***(372), pp.1351-1365.

Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q. and Nordborg, M., 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics,* ***44***(7), pp.825-830.

Serba, D.D., Daverdin, G., Bouton, J.H., Devos, K.M., Brummer, E.C. and Saha, M.C., 2015. Quantitative trait loci (QTL) underlying biomass yield and plant height in switchgrass. *Bioenergy Research*, ***8***(1), pp.307-324.

Shen, S., Li, X.F., Cullen, W.R., Weinfeld, M. and Le, X.C., 2013. Arsenic binding to proteins. *Chemical Reviews,* ***113***(10), pp.7769-7792.

Shin, H., Shin, H.S., Dewbre, G.R. and Harrison, M.J., 2004. Phosphate transport in *Arabidopsis*: Pht1; 1 and Pht1; 4 play a major role in phosphate acquisition from both low‐and high‐phosphate environments. *The Plant Journal,* ***39***(4), pp.629-642.

Shibghatallah, M.A.H., Khotimah, S.N., Suhandono, S., Viridi, S. and Kesuma, T., 2013. Measuring leaf chlorophyll concentration from its color: A way in monitoring environment change to plantations. *AIP Conference Proceedings,* ***1554*** (1), pp.210-213.

Silady, R.A., Effgen, S., Koornneef, M. and Reymond, M., 2011. Variation in seed dormancy quantitative trait loci in *Arabidopsis thaliana* originating from one site. *PloS One,* ***6***(6), p.e20886.

Simon, M., Loudet, O., Durand, S., Bérard, A., Brunel, D., Sennesal, F.X., Durand-Tardif, M., Pelletier, G. and Camilleri, C., 2008. Quantitative trait loci mapping in five new large recombinant inbred line populations of *Arabidopsis thaliana* genotyped with consensus single-nucleotide polymorphism markers. *Genetics,* ***178***(4), pp.2253-2264.

Singh, S., Mohapatra, T., Singh, R. and Hussain, Z., 2013. Mapping of QTLs for oil content and fatty acid composition in Indian mustard [*Brassica juncea* (L.) Czern. and Coss.]. *Journal of Plant Biochemistry and Biotechnology,* ***22***(1), pp.80-89.

Sinha, S., Sinam, G., Mishra, R.K. and Mallick, S., 2010. Metal accumulation, growth, antioxidants and oil yield of *Brassica juncea* L. exposed to different metals. *Ecotoxicology and Environmental Safety,* ***73***(6), pp.1352-1361.

Sonah, H., O'Donoughue, L., Cober, E., Rajcan, I. and Belzile, F., 2015. Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. *Plant Biotechnology Journal, **13***(2), pp.211-221.

Sondheimer, E. and Kertesz, Z.I., 1948. Anthocyanin pigments. A*nalytical Chemistry, **20***(3), pp.245-248.

Song, W.Y., Park, J., Mendoza-Cózatl, D.G., Suter-Grotemeyer, M., Shim, D., Hörtensteiner, S., Geisler, M., Weder, B., Rea, P.A., Rentsch, D. and Schroeder, J.I., 2010. Arsenic tolerance in *Arabidopsis* is mediated by two ABCC-type phytochelatin transporters. *Proceedings of the National Academy of Sciences, **107***(49), pp.21187-21192.

Stoeva, N., Berova, M. and Zlatev, Z., 2003. Physiological response of maize to arsenic contamination. *Biologia Plantarum, **47***(3), pp.449-452.

Storm, N. and Darnhofer-Patel, B., 2003. MALDI-TOF mass spectrometry-based SNP genotyping. *Single Nucleotide Polymorphisms: Methods and Protocols,* pp.241-262.

Tanksley, S.D., 2004. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *The Plant Cell, **16***(1), pp.181-189.

Tattini, M., Landi, M., Brunetti, C., Giordano, C., Remorini, D., Gould, K.S. and Guidi, L., 2014. Epidermal coumaroyl anthocyanins protect sweet basil against excess light stress: multiple consequences of light attenuation. *Physiologia Plantarum, **152***(3), pp. 585-598.

Tripathi, R.D., Srivastava, S., Mishra, S., Singh, N., Tuli, R., Gupta, D.K. and Maathuis, F.J., 2007. Arsenic hazards: strategies for tolerance and remediation by plants. *Trends in Biotechnology, **25***(4), pp.158-165.

Tu, C. and Ma, L.Q., 2002. Effects of arsenic concentrations and forms on arsenic uptake by the hyperaccumulator ladder brake. *Journal of Environmental Quality, **31***(2), pp.641-647.

Ullrich-Eberius, C.I., Sanz, A. and Novacky, A.J., 1989. Evaluation of arsenate-and vanadate-associated changes of electrical membrane potential and phosphate transport in *Lemna gibba* G1. *Journal of Experimental Botany*, pp.119-128.

van Eeuwijk, F.A., Bink, M.C., Chenu, K. and Chapman, S.C., 2010. Detection and use of QTL for complex traits in multiple environments. *Current Opinion in Plant Biology*, ***13***(2), pp.193-205.

Vatamaniuk, O.K., Mari, S., Lu, Y.P. and Rea, P.A., 1999. AtPCS1, a phytochelatin synthase from *Arabidopsis*: isolation and in vitro reconstitution. *Proceedings of the National Academy of Sciences,* U.S.A. ***96***(12), pp.7110-7115.

Vatamaniuk, O.K., Mari, S., Lu, Y.P. and Rea, P.A., 2000. Mechanism of heavy metal ion activation of phytochelatin (PC) synthase blocked thiols are sufficient for PC synthase-catalyzed transpeptidation of glutathione and related thiol peptides. J*ournal of Biological Chemistry, **275***(40), pp.31451-31459.

Wang, W.Y., Barratt, B.J., Clayton, D.G. and Todd, J.A., 2005. Genome-wide association studies: theoretical and practical concerns. *Nature Reviews Genetics, **6***(2), pp.109-118.

Wang, X., Peng, B., Tan, C., Ma, L. and Rathinasabapathi, B., 2015. Recent advances in arsenic bioavailability, transport, and speciation in rice. *Environmental Science and Pollution Research, **22***(8), pp.5742-5750.

Wang, Y., Hou, Y., Gu, H., Kang, D., Chen, Z.L., Liu, J. and Qu, L.J., 2013. The Arabidopsis Anaphase-Promoting Complex/Cyclosome subunit 1 is critical for both female gametogenesis and embryogenesis F. *Journal of Integrative Plant Biology*, ***55***(1), pp.64-74.

Wang, Z., Cheng, J., Chen, Z., Huang, J., Bao, Y., Wang, J. and Zhang, H., 2012. Identification of QTLs with main, epistatic and QTL× environment interaction effects for salt tolerance in rice seedlings under different salinity conditions. *Theoretical and Applied Genetics, **125***(4), pp.807-815.

Wilson, I.W., Schiff, C.L., Hughes, D.E. and Somerville, S.C., 2001. Quantitative trait loci analysis of powdery mildew disease resistance in the *Arabidopsis thaliana* accession Kashmir-1. *Genetics, **158***(3), pp.1301-1309.

Wojas, S., Clemens, S., SkŁodowska, A. and Antosiewicz, D.M., 2010. Arsenic response of AtPCS1-and CePCS-expressing plants–Effects of external As (V) concentration on As-accumulation pattern and NPT metabolism. *Journal of Plant Physiology, **167***(3), pp.169-175.

Wu, S., Fu, G., Chen, Y., Wang, Z. and Wu, R., 2012. Genetic mapping of complex traits by minimizing integrated square errors. *BMC genetics, **13***(1), pp.101-103.

Yang, Z. and Chu, C., 2011. Towards understanding plant response to heavy metal stress. *Abiotic Stress in Plants–Mechanisms and Adaptations*, pp.59-78.

Yang, H.C., Fu, H.L., Lin, Y.F. and Rosen, B.P., 2012. Pathways of arsenic uptake and efflux. *Current Topics in Membranes, **69***, p.325.

Zhang, J., Zhu, Y.G., Zeng, D.L., Cheng, W.D., Qian, Q. and Duan, G.L., 2008. Mapping quantitative trait loci associated with arsenic accumulation in rice (Oryza sativa). *New Phytologist, **177***(2), pp.350-356.

Zhao, F.J., McGrath, S.P. and Meharg, A.A., 2010. Arsenic as a food chain contaminant: mechanisms of plant uptake and metabolism and mitigation strategies. *Annual Review of Plant Biology, **61***, pp.535-559.

Zhu, J.J., Li, Y.R. and Liao, J.X., 2013. Involvement of anthocyanins in the resistance to chilling-induced oxidative stress in *Saccharum officinarum* L. leaves. *Plant Physiology and Biochemistry, **73***, pp.427-433.

Supplementary data 1

Recipe of Hoagland's complete nutrient solution:

Prepare the following stock solutions (1-6) for 1L working nutrient solution.

(1) 1.00 M $NH_4H_2PO_4$ use 1 mL/L of nutrient solution

(2) 1.00 M KNO_3 use 6 mL/L of nutrient solution

(3) 1.00 M $Ca(NO_3)_2$ use 4 mL/L of nutrient solution

(4) 1.00 M MgSO4 use 2 mL/L of nutrient solution

(5) Micronutrient stocks: combine the following amount of salts in a total volume of one liter of water, and then use 1 mL/L of this entire stock mixture (5) along with the stocks above (1-4) and the iron stock (6) described below to make up a total of 1 L of nutrient solution.

2.86 gm H_3BO_3

1.81 gm $MnCl_2 \cdot 4H_2O$

0.22 gm $ZnSO_4 \cdot 7H_2O$

0.08 gm $CuSO_4 \cdot 5H_2O$

0.02 gm $H_2MoO_4 \cdot H_2O$

(6) Iron stock: 40μM Fe-EDTA to the nutrient solution for final volume of 1L of working nutrient solution and store in a bottle covered with foil (dark).

Supplementary data 2

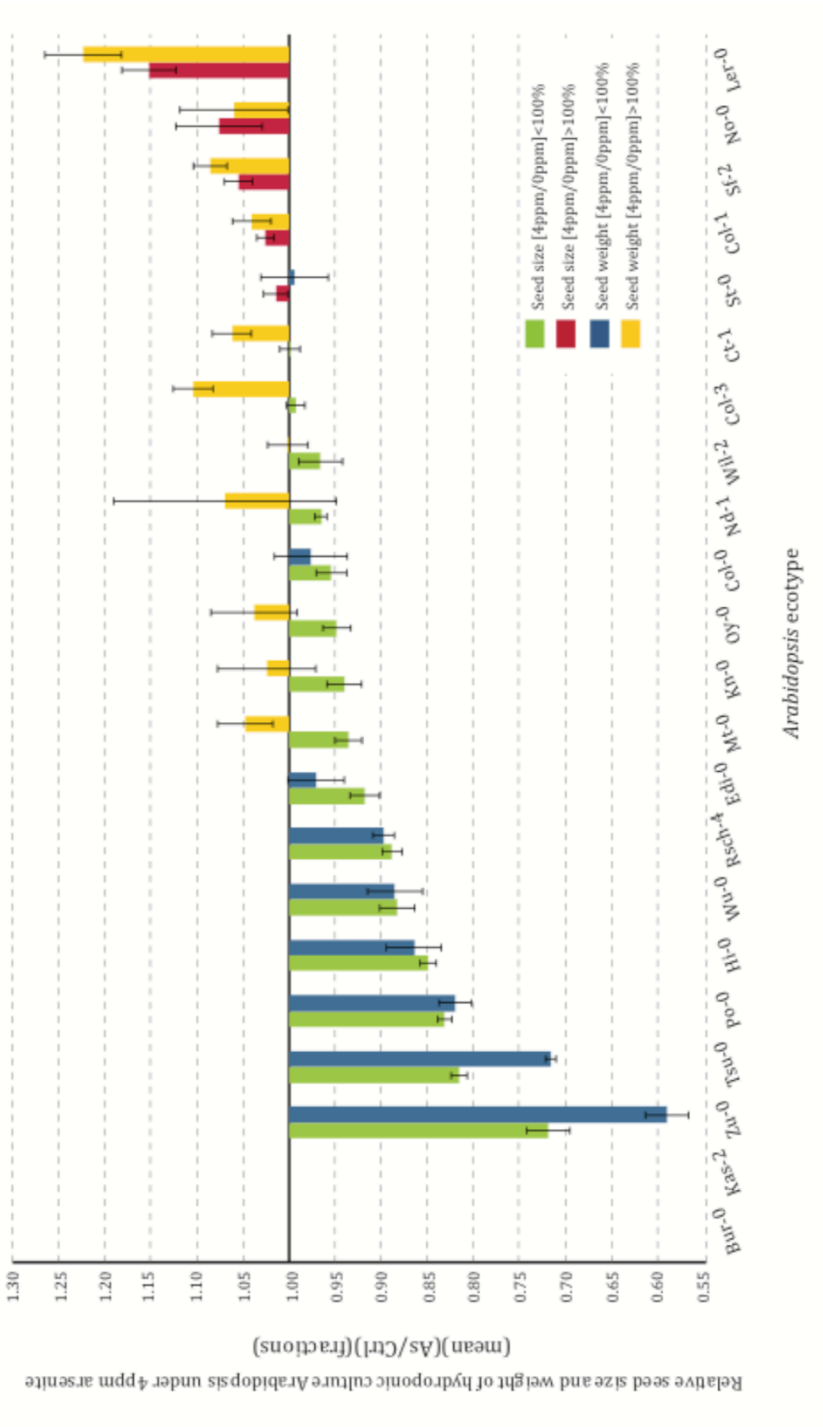

Figure 7.1. The mean relative seed size and weight of 22 *A. thaliana* ecotypes were grown under 4ppm sodium arsenite. Data are means of four replicates (±SD).

In Figure 7.1, X axis and Y axis cross at value 1.0. The values below 1 reflect the reduction in seed mass in the presence of arsenite and that smaller values indicate a stronger effect. Those sees from arsenite treated groups, which are larger than the control, are represented above the X-axis, and vice versa. Apart from genotypes Bur-0 and Kas-2, which did not produce seeds in this condition, genotype Zu-0 has the lowest value of both relative seed weight and size. Genotype Ler-0 has the highest value of both relative seed weight and size. Analysing genotypes Mt-0, Kn-0, Oy-0, Nd-1, Wil-2 and Col-3, the seed sizes from the arsenite treated group are smaller than the control, however the weights are heavier.

In general, the genetic background had a significant influence on the seed size (Two-way ANOVA, *F-value*=57.057, *P-value*=0.000) and weight (Two-way ANOVA, *F-value* =53.006, *P-value* = 0.000) of *A. thaliana* that was cultured hydroponically (Figure 3.19). The sodium arsenite treatment also had a significant effect on the seed size (Two-way ANOVA, *F-value*=256.206, *P-value*=0.000) and seed weight (Two-way ANOVA, *F-value*=253.820, *P-value*=0.000). The interaction between ecotypes and arsenite stress had a significant effect on seed size (Two-way ANOVA, *F-value*=60.648, *P-value*=0.000) and seed weight (Two-way ANOVA, *F-value*=94.179, *P-value*=0.000).

Supplementary data 3

	Marker																																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Segregation (A x B)																																							
1	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
2	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
3	B	-	-	B	A	A	B	-		B	B	-	-	-	B	-	B	B	B	B	-	-	B	-		B	B	-	B		B	B	-	B	B	B	B	B	A	A
4	-	A	A	A	A	A	A	A	A	A	-	A	A	A	A	A	-	-	-	-	A	A	A	A	A	A	-	A	A	A	A	A	A	A	A	A	A	A	A	A
5	B	-	-	B		A	-	-	-	-	B	-	-	-	-	-	B	B	B	B	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	B	-	A	A
6	A	A	A	A	B	A	A	A	A	A	A	A	A	A		A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
7	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
8	B	-	-	B		A	-	-	-	-	B	-	-	-	-	-	B	B	B	B	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	B	-	A	A
9	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	-
10	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B
11	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	B
12	A	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A

	Marker																																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Segregation (A x B)																																							
13	A	-	B	B	B	A	-	-	-	B	B	-	B	-	-	-	B	B	B	B	-	-	B	B	-	-	B	-	-	-	-	B	B	-	-	-	B	-	A	A
14	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
15	A	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	-
16	A	-	-	B	-	A	-	-	B	B	B	-	-	-	-	-	B	B	A	A	A	A	A	A	A	A		A	A	A	A	A	A	A	A	A	A	A	A	A
17	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
18	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	-	B	B	B	B	B	B		B	B	B	B		B	B	B	-	B	B	B
19	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-	A	A	A
20	-	B	B	B	B	A	B	B	B	B	-	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	B
21	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
22	B	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	B	A	B	B	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	A	A
23	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	A	A	A	A
24	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-	A
25	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
26	-	B	B	B	A	A	B	B	B	B	-	B	B	B	B	B	A	-	-	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
27	B	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-	-

	Marker																																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Segregation (A x B)																																							
43	-	A	A	A	A	A	A	A	A	A	-	A	A	A	A	A	-	-	-	-	A	A		A	A	A	A	A	A	A		A	A	A	A	A	A	A	B	B
44	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A
45	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
46	-	A	A	A	B	A	A	A	A	A	-	A	A	A	A	A	B	-	-	-	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A
47	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
48	A	-	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
49	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
50	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
51	A	B	B	B	B	A	B	B	B	B	-	B	B	B	B	B	B	-	-	-	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
52	A	A	A	A	-	A	A	A	A	A	A	A	A	A		A	A	A	A	A	-	-	-	-	-	-	-	-	-	-	-	A	-	B	B	B	A	-	B	B
53	-	A	A	A	B	A	A	A	A	A	-	A		A	A	A	A	-	-	-	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
54	A	-	-	A	A	A	-	A	A	-	A	-	-	-	-	-	A	A	A	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
55	B	B	B	B	A	A	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B
56	B	B	B	B	A	A	B	B	B	B	B	B		B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
57	A	A	A	A		A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	-	-	B	B	B	B	B	B	B	B	B	B	-	B	B	B	B	B	A	A

	Marker																																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Segregation (A x B)																																							
58	A	A	A	A	B	A	B	B	B	B	A	B	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
59	B	-	-	B	A	A	-	-	-	-	B	-	-	-	-	-	B	B	B	B	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	B	-	B	B
60	A	B	B	B	B	A	B	B	B	B	A	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
61	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
62	B	B	B	-	A	A	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	A	B	A	B	B
63	A	A	A	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
64	A	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
65	A	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
66	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	-	A	A	A	A	A	A	A	A	B	B	A	A
67	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
68	B	B	B	A	B	A	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
69	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
70	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
71	B	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
72	-	A	A	-	A	A	A	A	A	A	-	A	A	A	A	A	-	-	-	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-	A	B	B

	Marker																																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Segregation (A x B)																																							
73	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-
74	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
75	-	B	B	B	B	A	B	B	B	B	-	B	B	B	B	B	B	-	-	-	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
76	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	-	-
77	A	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	-	B	-		B	B	-	B	B	B	B	B	B	B	A	-	A	A	A
78	B	B	B	B	A	A	B	B	B	B	B	-	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
79	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
80	B	-	-	B	-	A	B	-	-	B	B	-	-	-	B	-	B	B	B	B	-	-	B	-	-	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
81	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
82	B	B	B	B	A	A	B	B	B	B	B	B		B	B	B	B	B	A	A	A	A	A	A	A	A	-	A	A	A	A	A	A	A	A	A	A	A	A	A
83	A	A	A	-	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B
84	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
85	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
86	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
87	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B

	Marker																																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Segregation (A x B)																																							
88	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	-	A	A	A	A	A	A	A	A	A	A	A	A	B	B
89	A	B	B	A	-	A	B	B	B	B	A		B	B	B	B	A	A	A	A		B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	A	A
90	A	A	A	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
91	-	A	A	A	A	A	A	A	A	A	-	A	A	A	A	A	A	-	-	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
92	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-	B
93	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
94	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
95	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	-
96	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-		B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
97	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	B
98	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A
99	A	A	A	-	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
100	B	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	-
101	B	A	A	B	A	A	A	A	A	A	B	A	A	A	A	A	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	-	-
102	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A

	Marker																																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Segregation (A x B)																																							
103	B	B	B	B	B	A	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	A
104	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B